

OM of: US-09-303-518d-125 to: PIR-71.* out_format: pfs
 Date: Jun 30, 2002 7:39 AM

About: Results were produced by the GenCore software, version 4.5,
 Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODE=framer_gdp_model -DEV=x1h
 -CGEN2.1/USPRO-us09303518/runat_28062002_142713.4317/app_query.fasta.1.23501
 -DB-PIR-71 -QFMT=fastan -SUFFIX=trp -GAPOP=12.000 -GAPEXT=4.000
 -MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000 -OGAPOP=4.500
 -OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
 -FGAPEXT=7.000 -YGAPEXT=10.000 -YGAPEXT=0.500 -DELTOP=6.000
 -DELEXT=7.000 -START=1 -MATRIX=blomsu62 -TRANS=human40.cdi
 -LIST=100 -DOCAIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
 -ALIGN=45 -MODE=LOCAL -OUTFMT=pfs -NORM-ext -HEAPSIZE=500
 -MCPEN=0 -MAXLEN=200000000 -USER=US09303518.@CGEN1.1.1092
 -NO_XUPXY -WAIT -THREADS=1

Search information block:

Query: US-09-303-518d-125
 Query length: 1344
 Database: PIR-71.*
 Database sequences: 283138
 Database length: 96089334
 Search time (sec): 504.980000

score list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
PIR2:D81185	2289.00	3158.37	1.8e-168	447	sodium-translocating NADH dehyd
PIR2:A61919	2251.00	3105.74	1.5e-165	447	probable sodium-translocating N
PIR2:144002	1629.00	2244.24	1.5e-117	447	sodium-translocating NADH dehyd
PIR2:A60393	1596.00	2198.54	5.1e-115	447	NADH dehydrogenase (ubiquinone)
PIR2:551015	1468.50	2021.97	3.5e-105	446	sodium-translocating NADH dehyd
PIR2:682094	1455.50	2003.60	3.6e-104	446	sodium-translocating NADH dehyd
PIR2:HB3272	1353.00	1862.02	2.8e-96	445	sodium-translocating NADH dehyd
PIR2:CG1751	477.00	639.96	3.2e-28	465	probable sodium-translocating N
PIR2:DB6583	464.50	630.92	1.0e-27	467	ubiquinone oxidoreductase, aliph
PIR2:ET2040	464.00	605.34	2.7e-26	465	probable sodium-translocating N
PIR2:F71489	158.00	200.91	0.0005	801	sodium-translocating NADH dehyd
PIR2:282252	150.00	194.25	0.0019	519	mtfc protein - Rhodospirillum rubrum
PIR2:AE0692	144.50	183.98	0.0055	673	probable NADH reducing dehydrog
PIR2:ET2398	140.00	181.82	0.0109	451	hypothetical protein TM0244 - T
PIR2:SE6136	138.50	173.68	0.0170	819	mtfc protein homolog - Haemophi
PIR2:SE69703	135.00	160.81	0.0403	1802	HKRI protein precursor - Yeast
PIR2:TO2345	131.00	155.33	0.0818	1791	hypothetical protein KIA00324
PIR2:DA0670	129.50	157.33	0.0945	1199	nuclear envelope protein POM 1
PIR2:DB8208	129.00	161.09	0.0904	774	probable ferredoxin PA3491 [Imp
PIR2:AA4943	126.00	161.95	0.1326	473	membrane protein [Imported] - E
PIR2:TS3903	125.00	160.71	0.1375	534	serine-rich protein - fission y
PIR2:SO1359	125.00	164.29	0.1416	328	salivary glue protein sgs-3 pr
PIR2:TS5310	124.00	150.56	0.2448	1104	probable membrane protein YMR3
PIR2:TS5352	118.00	144.82	0.6580	857	hypothetical protein P884.120 -
PIR2:G66915	115.00	142.16	1.07	740	probable iron-sulfur protein bl
PIR2:F8769	115.00	142.16	1.07	740	probable membrane protein [Imp
PIR2:B90921	115.00	141.73	1.09	580	probable mucin DKFZP434C196.1 -
PIR2:TS43481	114.50	143.95	1.09	528	gastric mucin (clone PGM-2A) -
PIR2:147141	113.50	143.05	1.26	819	pyruvate, water dikinase (EC 2.7
PIR2:F75196	113.00	139.05	1.44	437	hypothetical protein 3 (gyrb re
PIR2:CG3135	112.00	143.36	1.56	631	epistatin - mouse
PIR2:132257	112.00	139.63	1.74	376	adenine deaminase-related prote
PIR2:CV5580	111.50	144.20	1.62	529	hypothetical protein [Imported]
PIR2:TS4311	110.50	140.03	1.97	2073	fatty acyl-CoA synthase (EC 2.
PIR2:TS3207	110.50	125.45	3.26	2073	fatty acid synthase, subunit b
PIR2:TS2869	110.00	125.93	3.44	1847	hypothetical protein T2332.1 -
PIR2:H70108	109.50	134.10	2.89	773	hypothetical protein BB0072 - I
PIR2:TS3409	109.50	125.63	3.72	1777	hypothetical protein T19D12.1 -

PIR2:TS34434	109.50	123.31	3.99	2232	hypothetical protein R0649.
PIR2:J00405	108.50	129.07	3.84	1106	hypothetical 119.5K protein
PIR2:A98048	108.00	130.50	3.94	898	H ⁺ /K ⁺ -exchanging ATPase (EC
PIR2:A36904	107.00	135.78	3.86	466	prochlorophyllide reductase
PIR2:EB95180	107.00	128.93	4.74	914	cation-transporting ATPase.
PIR2:TS33369	107.00	125.54	5.24	1275	hypothetical protein H02809
PIR2:TS4433	107.00	127.00	5.37	1367	glucan 1,4-alpha-glucosidase
PIR2:TS2123	105.50	128.86	5.65	1032	hypothetical protein R0649.
PIR2:TS21476	105.50	128.00	5.97	680	probable endoglucanase - Rum
PIR2:AS5575	105.50	110.92	9.96	817	pyruvate, water dikinase (EC
PIR2:164184	105.00	133.21	5.47	4377	ankyrin 3, long splice form
PIR2:CG7587	105.00	129.57	6.10	457	UDP-N-acetylmuramoylalanine-D
PIR2:1PWP694	104.50	133.00	6.07	654	hypothetical protein CC2730
PIR2:AD1994	104.50	131.52	6.16	481	H ⁺ -transporting two-sector A
PIR2:BD1229	104.50	126.56	7.15	504	hypothetical protein al1505
PIR2:TS2545	104.50	119.72	8.77	821	pyruvate, water dikinase (EC
PIR2:TS45025	104.00	110.92	12.22	1609	probable membrane protein Y
PIR2:BD4055	103.50	133.45	6.66	3570	mucin MUC5B, tracheobronchyl
PIR2:TS22509	103.50	130.61	7.25	481	phospho-2-dehydro-3-deoxyhep
PIR2:CG5982	103.00	131.33	7.60	364	H ⁺ -transporting ATP synthase
PIR2:098209	103.00	129.67	7.98	419	UDP-N-glucosamine 1-carboxy
PIR2:AE3077	103.00	129.67	7.98	493	aldhyde dehydrogenase dhas
PIR2:EB3157	102.50	123.55	8.58	466	aldhyde dehydrogenase dhas
PIR2:AA4917	102.50	124.32	10.03	779	hypothetical protein PA3909
PIR2:AE3077	102.00	129.94	9.08	419	UDP-N-acetylglucosamine 1-ca
PIR2:AE3077	102.00	122.06	11.84	1003	purine synthesis multifunct
PIR2:1P4578	102.00	119.08	12.57	1219	nucleoporin Nup153 homolog
PIR2:SB6813	101.50	128.50	10.14	451	hypothetical protein s11481
PIR2:SB6629	101.50	127.35	10.50	505	mucin 6, gastric (3-repeat C
PIR2:J04301	101.50	124.11	11.57	682	amphipysin - chicken
PIR2:1A1144	101.50	119.15	13.42	694	folliotropin receptor - pig
PIR2:1B64925	101.00	128.88	10.74	1131	hypothetical serine-rich pr
PIR2:1B84213	101.00	123.62	11.84	406	amino transferase nifs homolo
PIR2:670704	101.00	123.38	12.66	559	hypothetical protein Vng0553
PIR2:684693	101.00	120.88	13.33	697	hypothetical protein Rv3226C
PIR2:582525	101.00	116.99	15.33	891	probable proline-rich protei
PIR2:682433	100.50	130.51	10.94	1306	MSR2 protein - yeast (Sach
PIR2:10FBSHS	100.50	128.21	11.72	323	fructokinase VCA0656 [Import
PIR2:10FBSHS	100.50	123.90	12.56	405	ceoa protein - Burkholderia
PIR2:10FBSHS	100.50	122.29	14.00	508	histidine ammonia-lyase (EC
PIR2:10FBSHS	100.50	111.28	19.47	2135	a-agglutinin core protein Ag
PIR2:444618	100.50	112.41	14.93	739	voltage-dependent calcium c
PIR2:147425	100.00	118.43	16.82	2321	noct3 protein - human
PIR2:147425	100.00	111.09	20.96	3507	hypothetical protein ZK783.
PIR2:147425	100.00	122.83	22.64	642	timbin - yeast (Saccharomy
PIR2:147425	100.00	112.41	14.93	669	hypothetical protein Y1092 -
PIR2:147425	100.00	118.43	16.82	989	insulin II gene enhancer-bin
PIR2:147425	100.00	111.09	20.96	2035	host cell factor C1 precurs
PIR2:147425	100.00	112.41	14.93	406	selenocysteine lyase [Import
PIR2:147425	99.50	126.80	14.01	406	selenocysteine lyase [Imila
PIR2:147425	99.50	122.33	16.02	630	tumor-associated mucin (MUC1
PIR2:147425	99.50	120.91	16.72	724	ATP-dependent protease [Im
PIR2:147425	99.50	119.94	17.21	797	glycoprotein X precursor - e
PIR2:147425	99.50	116.86	18.88	1079	gag-fml1-env polypeptide -
PIR2:147425	99.50	113.60	20.81	1486	glutamine synthase (NAPDH)
PIR2:147425	99.50	103.08	28.52	4180	hypothetical protein PA0690
PIR2:147425	99.00	123.43	15.59	528	conserved hypothetical prote
PIR2:147425	99.00	107.71	26.57	2476	zonadhesin - pig

seq_name: PIR2:D81185

seq_documentation_block:

sodium-translocating NADH dehydrogenase (ubiquinone) (EC 1.6.5.5.) alpha chain NMB0569

Cispecies: Neisseria meningitidis

Cispecies: Neisseria meningitidis

Cispecies: Neisseria meningitidis

Cispecies: Neisseria meningitidis

Cispecies: Neisseria meningitidis

Cispecies: Neisseria meningitidis

A:Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755
A:Accession: D81185
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1447 <TET>
A:Cross-references: GB:AE002412; GB:AE002098; NID:97225783; PIDN:AAF40997.1; PID:9722579
A:Experimental source: serogroup B, strain MC58
A:Gene: NMB0569
C:Superfamily: Vibrio alginolyticus sodium-translocating NADH dehydrogenase (ubiquinone)
C:Keywords: oxidoreductase

alignment_scores:
Quality: 2289.00 Length: 447
Ratio: 5.121 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-303-518d-125 x D81185 ..

Align seg 1/1 to: D81185 from: 1 to: 447

```
1 ATGATTAATAATGAAAAAGTCTAAACCTGCCCATCGCGGCAGACCGGA 50
|||||
1 Metilelysielysylglyleuasnleuprollealaglyargprggl 17
51 GCAACCGCTTACGACGCGCGCCATACGAGTCCGCTGGTGGCG 100
|||||
17 uGlnalaVallyrAspGlyProAlaIleThrGluValAlaLeuLeuGlyG 34
101 AAGAATATGCGGTATGCGCCCTCGATGAAGTCAAGGAAGCGATGCC 150
|||||
34 luGluTyrAlaGlyMetArgProSerMetLysValLysGluGlyAspAla 50
151 GTCAAAAAGGCCAGTCTGTTGAACAAAAAGAAATCCGGCGTGT 200
|||||
51 ValLysLysGlyGlnValLeuPheGlnAspLysAsnProGlyValVa 67
201 GTTTCCTGCGCGCTTCAGCGCAAAATCGCGCGATTCACCGTGGCGAA 250
|||||
67 lPheThrAlaProAlaSerGlyLysIleAlaIleHisArgIleGlu 84
251 AGCGGCTACTCAGTCACTGCTGATGCGCTGGAAGGCAACGCAATC 300
|||||
84 ysArgValLeuGlnSerValIleAlaValGluGlyAsnAspGluIle 100
301 GAGTTGAACGCTAGCAGCCTGGAAGCGTGGCAACTTAAGCGCGAAGA 350
|||||
101 GlnPheGlnArgTyrAlaProGlnAlaLeuAlaAsnLeuSerGlyGlu 117
351 AGTGGCGCGCAACCTGATCAATCGGTTTGGACTGCGTGGCGACCC 400
|||||
117 uValArgArgAsnLeuIleGlnSerGlyLeuThrPThrAlaLeuArgThr 134
401 GTCCGTTACGAAAAATTCCTCGCTGATGCCGAGCGCTTGGCATTTTC 450
|||||
134 rGProPheSerLysIleProAlaValAlaAspAlaGluProPheAlaIlePhe 150
451 GTCAATGCGATGAGACCAATCCGCTGGCTGGCGACCCCTACGCTATT 500
|||||
151 ValAsnAlaMetAspThrAsnProLeuAlaAlaAspProThrValIle 167
501 CAAGAAGCCGCGAGGATTCGAACGCGCGCTTGTGTTTGAAGCGGT 550
|||||
167 elySGluAlaIleGluAspPheLysArgGlyLeuLeuValLeuSerArgL 184
551 TGACCGAAGCGCAAAATCCATGTTTGAAGGACGCTGGCGAGAGAGTCCG 600
|||||
184 eutThrGluArgLysIleHisValLysLysAlaIleGlyAlaAspValPro 200
601 TCTGAAAATGCTGCCAAACATCGAAACACATGAAATTGCGGCGCGCATCC 650
```

```
|||||
201 SerGluAsnAlaIleAsnIleGluThrHisGluPheGlyGlyProHisPr 217
651 TGGCGTTTACGAGCGACGACATTCATTTATCATGACCGCGTGGCGCA 700
|||||
217 oAlaGlyLeuSerGlyThrHisIleHisPheIleGluProValGlyAla 234
701 ATAAACCGTGTGACCATCAATTCATCAAGATGATTAATTCATGGCCGT 750
|||||
234 snLysThrValIleThrIleAsnTyrGlnAspAlaIleThrIleGlyArg 250
751 TGTTCGCAACAGCGCGTTCGACACCGACCGCGTATTCCTAGTGG 800
|||||
251 LeuPheAlaThrGlyArgLeuAsnThrGluArgValIleAlaLeuGlyG 267
801 TTCTCAAGTCAACAACCGCGCTCTTGGTACCGCTTTGGGTCGCAAG 850
|||||
267 ySerGlnValAsnLysProArgLeuLeuArgThrValLeuGlyAlaLysV 284
851 TATCGCAATTAATCTGCGCGGATTCGTTGACACAGACCGCGTATTT 900
|||||
284 alSerGlnIleThrAlaGlyGluLeuValAspThrAspAsnArgValIle 300
901 TCCGTTTCGCTATTTGAACGCGGATTCACACAGCGCGCATTAATT 950
|||||
301 SerGlySerValLeuAsnGlyAlaIleThrGlnGlyAlaHisAspTyr 317
951 GGGACGCTACCAATCAATCAATTCGCTTACGAGAGCGCGCAGCAAG 1000
|||||
317 uGlyArgTyrHisAsnGlnIleSerValIleGluGluGlyArgSerLysG 334
1001 AGCTGTTTCGCTGCTTGGCGCGCGACCGACAAATATCTCATCAGCGT 1050
|||||
334 lueupPheGlyTyrValAlaProGlnProAspLysTyrSerIleThrArg 350
1051 ACAACCTCGGCGATTTCTGAAAAACAACCTTCAAGTTCAACACAGC 1100
|||||
351 ThrThrLeuGlyHisPheLeuLysAsnLysLeuPheLysPheAsnThrAl 367
1101 CGTCAACGCGCGCGCGCGCGCGATTCGCGGATTCGATTCAGCGCG 1150
|||||
367 aValAsnGlyGlyAspArgAlaMetValProIleGlyThrTyrGluArgV 384
1151 TGATGCCCTTGGATATTCCTGCCACCGCTCTTGGCGATTTAATCGTC 1200
|||||
384 alMetProLeuAspIleLeuProThrLeuLeuLeuArgAspLeuIleVal 400
1201 GCGGATACGACAGCGCGCGCATTCGCTGCTTGAATTCGACGAGA 1250
|||||
401 GlyAspThrAspSerAlaGlnAlaLeuGlyCysLeuGluLeuAspGlu 417
1251 AGACCTCGCTTGTGACGCTGCTGCGCGCGCAATTCGAATTCAGCGC 1300
|||||
417 unsPleuAlaLeuCysSerPheValCysProGlyLysTyrGluTyrGlyP 434
1301 CGCTGTTGCGCAAAAGTGTGAACCATTCAGAGAAGAGGC 1341
|||||
434 roLeuLeuArgLysValLeuGluThrIleGluLysGluGly 447
```

seq_name: p1r2:A81919

seq_documentation_block:

probable sodium-translocating NADH dehydrogenase (ubiquinone) (EC 1.6.5.-) chain A NM

C:Species: *Neisseria meningitidis*
C:Date: 05-May-2000 #sequence-revision 05-May-2000 #text-change 16-Feb-2001

C:Accession: A81919
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo

; Holtroyd, S.; Jorgensen, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre

Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491

A:Reference number: A81775; MUID:20222556
A:Accession: A81919
A:Status: preliminary

A: Molecule type: DNA
 A: Residues: 1-447 <PAR>
 A: Cross-references: GB:AL162754; GB:AL157959; NID:g7379424; PIDN:CAB84035.1; PID:g737947
 A: Experimental source: serogroup A, strain Z2491
 C: Genetics:
 A: Gene: ngrA; NMA0752
 C: Superfamily: Vibrio alginolyticus sodium-translocating NADH dehydrogenase (ubiquinone)
 C: Keywords: NAD; oxidoreductase

alignment_scores:
 Quality: 2251.00 Length: 447
 Ratio: 5.036 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 97.763

alignment_block:

US-09-303-518D-125 x A81919 ..

Align seg 1/1 to: A81919 from: 1 to: 447

```

1 ATGATTAATAAATAAAGCTTAACCTGCGCATCGCGGACAGCGGA 50
|||||
1 Metilelystlelysglyleuasnleuprollealaglyargprogl 17
51 GCAAGCCGTTACGACGGCCGCGCATTCAGCAAGTCGCGTGGCG 100
|||||
17 uGlnValIleIetyraspglyProValIleThrglValAlaIleuenglyg 34
101 AAGAAATATGCGGTATGCGCCCTCGATGAAGTCAGGAAGGCGATGCC 150
|||||
34 lUGlUtyAlaIglMetArpGosermetlyVallysglUglYAspAla 50
151 GTCAAAAAGCGCAGTGTGTTGAAGCAAAAAGATCCGGCGCTGT 200
|||||
51 VallysglysglylnValleuphegluasplysasnProglYValIva 67
201 GTTACTGCGCGCGCTTACGCAAAATCGCGCGATTCACCGTGGCGAA 250
|||||
67 lPheThrAlaProValIserglylyslleAlaIleAlaIleHisargglyul 84
251 AGCGGCTTACGTCAGTGTGATGCGCGTGAAGGACAGCAAGATC 300
|||||
84 ysargValleuInserValIleAlaIleAlaIleAlaIleAsnAspIle 100
301 GAGTTGAAGCTACGACCTGAAGCGCGCAAACTTAAGCGGGAAGA 350
|||||
101 GluphegluaIrglyAlaProglAlaIleuAlaIleuSerglyglUgl 117
351 AGTGGCGCGCAACCTGATCAATCGGTTGTGACTGCGCTGCGCACCC 400
|||||
117 uValArgrAlaIleuIleInserglyleuThrAlaIleuArgrThra 134
401 GTCCGTTACGCAAAATTCGCGCGTGCATCGCGCGTGGCGATCTTC 450
|||||
134 rgrProPheSerlyslleProAlaValaIleAlaIleProPheAlaIlePhe 150
451 GTCAATGCGATGACCAATCGCGTGGCGCGACCTCAAGGATCATAT 500
|||||
151 ValaIleAlaIleMetAspThrAsnProleuAlaIleAspProValIle 167
501 CAAGAAGCGCGCGAGATTCGAAGCGCGCGCTGTGATTAAGCGCGT 550
|||||
167 elysglUlnAlaIleaglUasPheArgrArgglyleuIleuValleuSerargl 184
551 TACCGAAGCGCAAAATTCATGTTTGAAGCGCGCTGGCGCGACAGTCGCC 600
|||||
184 eUrThcluarIrglyslleHisValIleValaIleAlaIleAlaIleAspValPro 200
601 TGTGAAAATGTCGCAACATCGAAGCAACATGAATCGCGCGCGCGCATCC 650
|||||
201 SergluAsnAlaIleAsnIleglUthrHsglUupheglYglYProHisPrt 217
651 TCCCGGTTTGAAGTGCAGCGACATTCATTTCATCGAGCGCGTGGCGGA 700

```

```

|||||
217 calaglyleuSerlyThrHisIleHisPheIleglUuproValIglYAla 234
701 ATAAAACCGGTGGACCATCATTAATCAAGATGTAATTCATGGCGCT 750
|||||
234 snlystrValIleThrIleAsnIleAsnIleAsnIleAlaIleAlaIle 250
751 TTTTGGCAACAGCGCGTGTGAACAGCGAGCGGTGATTCCTAGTGG 800
|||||
251 leuPheAlaThrIleIleValleuAsnIleThrIleValIleAlaIle 267
801 TTTCAAGTCACAAACCGCGCTTGGTGTACCGTTTGGTGGCGAAG 850
|||||
267 ySerGlInValaIleAsnIleProArgrleuIleuArgrThrValleu 284
851 TATCGCAATTAATCTGCGCGCAATTTGTTGACACAGCAACCGCGT 900
|||||
284 alSerGlInIleThrAlaIleaglUleuValaIleAspAlaIleAla 300
901 TCCGGTTGCGTATTAAGCGCGCATTCACAAAGCGCGCGACGATTA 950
|||||
301 SerGlySerValleuAsnIleAlaIleThrglInglYAlaIleAsp 317
951 GGGACGCTACACACATCAGATTCGTTATCAAGAGCGCGCGCAAG 1000
|||||
317 uGlyArgTyHisAsnIleIleSerValIleglUglUglYArgSerlysg 334
1001 AGCTGTTCGCGTGGTGGCGCGCGCGACCGGCAAAATCTCATCAGCG 1050
|||||
334 lUeupheglUtyrValAlaIleProglInProAspIlystrIleThrArg 350
1051 ACAACCGTGGCGCATTCCTGAAAACAAACCTTCAGTCAACAGCAC 1100
|||||
351 ThrThrIleuGlYHisPheIleuIleAsnIleuIlePheIleThrAla 367
1101 CGTCAACGGCGCGGACCGCGCATTCGATTCGATTCAGACGCGG 1150
|||||
367 aValaIleAsnIleglYAspArgAlaMetValProIleglYThrTyglUarg 384
1151 TGATGCCCTTGGATATTCCTCCACCGCTGTTGGCGCATTAATCGTC 1200
|||||
384 alMetProleuAspIleIleuProThrIleuIleuIleuIleuIleVal 400
1201 GCGCATACCGACAGCGCGCGCATTCGCTTGGATTCGACGCAAGA 1250
|||||
401 GlyAspThrAspSerAlaIleAlaIleuIleCysleuIleuAspIle 417
1251 AGACCTTCGCTTGTGACGCTTCGCTGCGCGGCAAAATCAGATACGCC 1300
|||||
417 uAspIleuAlaIleuIleuSerPheValIleCysProglYIleTyglUtyglY 434
1301 CGCGTGGCGCAAAAGTGTGAAACCAATTCGAAAGGAGGCG 1341
|||||
434 rUleuIleuArgrlyValleuIleuThrIleglUlysglyUgl 447

```

seq_name: p1r2:164002

seq_documentation_block:

sodium-translocating NADH dehydrogenase (ubiquinone) (EC 1.6.5.-) alpha chain H10164
 C:Species: Haemophilus influenzae
 C:Date: 18-Aug-1995 #sequence_revision 02-Sep-2000 #text_change 02-Mar-2001
 C:Accession: 164002; A64003
 R:Flaischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage,
 ; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodex, A.; Kelley, J.M.; Weidman,
 ; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.;
 Science 269, 496-512, 1995
 A:Authors: Grehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Vente
 A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
 A:Reference number: A64000; MUID:95350630
 A:Accession: 164002
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-447 <TIGR>

A:Cross-references: GB:I42023; TIGR:H10164; GB:U32702; NID:g1573118; PIDN:AAC21836.1; P1
 A:Note: the sequence is revised in Genbank entry U32702, PID:g1573122
 C:Superfamily: Vibrio alginolyticus sodium-translocating NADH dehydrogenase (ubiquinone)
 C:Keywords: oxidoreductase

alignment_scores:
 Quality: 1629.00 Length: 448
 Ratio: 4.093 Gaps: 2
 Percent Similarity: 88.839 Percent Identity: 68.973

alignment_block:
 US-09-303-518d-125 x I64002 ..

Align seg 1/1 to: I64002 from: 1 to: 447

```

1 ATGATTAATAATCAAAAAGGTCTAAGCCGATCGGCGGAGACGCGGA 50
1 MetIleThrIleLysLysGlyLeuAspLeuProIleAlaGlyLysProAl 17
51 GCAGCCCGTTTACGACGGCGCCGATACCGAAGTCCGTTGCTGGCG 100
17 aGlnValIleHisSerGlyAsnAlaValAsnGlnValAlaIleLeuGly 34
101 AAGATATACCGCGTATGCGCCCTCGATGAAAGTCAAGGAAGCGATCC 150
34 IuGluIuValGluMetArgProSerMetLysValArgGluGlyAspVal 50
151 GTCAAAAAGGCGCAAGTCTGTTGAACAAAAAAGATCGGCGCGTGT 200
51 ValLysLysGlyGlnValIleLeuPheGlnLysPylAsnProGlyValAl 67
201 GTTACTGCGCGCGCTTCAGGCAAAATCGCGCGATTCACCGTGGCGAA 250
67 ePheThrAlaProAlaSerGlyThrIleThrAlaIleAsnArgGlyL 84
251 AGCGGCTACTTCAGTCACTGCTGATTCGCGTTGAAGGACGACGAAATC 300
84 YsArgValLeuGlnSerValIleAsnValGluGlyAspGluLysIle 100
301 GAGTTTGAAGCGTACGACGCTGAAGCGCTGCGCAACTTAACGGCGAA 350
101 ThrPheAlaLysTyrSerThrGluGlnLeuAsnThrLeuSerSerGlu 117
351 AGTGGCGCGCAACTGATCAATCCGCTTGTGAGTGCCTGCGCAACC 400
117 nValLysGlnAsnLeuIleGlnSerGlyLeuTrpThrAlaLeuArgH 134
401 GTCCGTTAGCAAAATTCCTCCGTCGATGCGCGGACCGCTTCGCATC 450
134 rGpProPheSerLysValProSerIleGlnSerGluAlaSerSerIle 150
451 GTCAATGCGATGACACCAATCCGCTGCTGCGACCTACGCGCATAT 500
151 ValAsnAlaMetAspThrAsnProLeuAlaAspProSerValValle 167
501 CAAGAAGCCGCGAGATTTCAACGCGCGCTGTTGATTTGAGCCGTT 550
167 uLysGluLysSerGlnAspPheThrAsnGlyLeuThrValLeuSerArg 184
551 TGACCGAAGCAAA...ATCATGTTTGTAAAGGACGCTGCGCGACGCTG 597
184 eUpheProSerLysProLeuIleLysLysValGlyAspSerAsnIle 200
598 CCGTCTGAAATGCTGCCAATCATGAAACATGATTTGCGCGCGCGCA 647
201 ProThrAlaAspLeuGlnAsnLeuGlnIleHisAspPheThrGlyVal 217
648 TCGTCCGCGTTTGAAGTGCACGACATTTTCATGAGCGCGTGGCG 697
217 sProAlaGlyLeuValGlyThrHisIleHisPheIleAspProValGly 234
698 CGAATAAACCCTGTGACCATCATATTATCAAGATGTAATTACCATTC 747

```

```

234 legIuLysThrValThrPheHisIleAsnTyrGlnAspValIleAlaValGly 250
748 CGTTTGTTCGAACAGCGCGCTGCAACACCGACCGCGTATTCCTCCTAG 797
251 LysLeuPheThrThrGlyGlnLeuTyrSerGluArgValIleSerLeuAl 267
798 TGGTTCTCAAGTCAACAAACCGCGCTGCTGACCGTTTGGTGGCGA 847
267 aGlyProGlnValLysGluProArgLeuValAlaGlyThrIleGlyAla 284
848 AAGTATGCAAAATTTACTGCGCGCAATGGTTGACACAGCAACCGCGT 897
284 snLeuSerGlnLeuThrGlnAsnGlnLeuSerAlaGlyAsnArgVal 300
898 ATTTCCGCTTCGATTTGAACGCGCGATTTACCAAGCGCGCGACATTA 947
301 IleSerLysSerValLeuLysGlnIleAlaLysAspSerHisAsp 317
948 TTTGGACGCTTACCAATGAGATTTCCGTTATCGAAGAGCGCGACGA 997
317 rLeuGlyArgTyrAlaLeuGlnValSerValIleAlaGluGlyAsnGlu 334
998 AAGAGCTGTTCGCGTGGTTCGCGCGACCGCGCAAAATTCATCATCAG 1047
334 YsGluPhePheGlyTyrIleMetProGlnAlaAsnLysTyrSerValThr 350
1048 CGTCAAAACCTCGCGCATTTCTGAAAAAACAATCTTCAGTTCAACAC 1097
351 ArgThrValLeuGlnIleHisPheSerLys...LysLeuPheAsnThr 366
1098 AGCGCTCAACGCGCGCGACCGCGCATGCGCGATTTGATCTACGACG 1147
366 rSerGlnAsnGlyGlyGlnArgAlaMetValProIleGlySerTyrGlu 383
1148 GCGTGAATGCCCTTGATATCTGCGCCACCGCTTGTGCGCATTAATC 1197
383 rValMetProLeuAspIleLeuProThrLeuLeuArgLysPheLeu 399
1198 GTGCGGATACGACGACGCGCGACGATTTGGTGTGATTTGGACGA 1247
400 ValGlyAspThrAspGlyAlaGlnGlnLeuGlyLysLeuGlnLeuAsp 416
1248 AGAAGACCTCGCTTGTGACGCTGCTGCGCGCGCAATGCAATGACG 1297
416 uGlnAspLeuAlaLeuLysSerPheValCysProGlyLysTyrGluTyr 433
1298 GCCCGCTGTTCGCAAAAGTGTGAAACCATTTGAGAGGAGAGC 1341
433 YsSerIleLeuArgGlnValLeuAspLysIleGluLysGluGly 447

```

seq_name: p1r2:AG0393

seq_documentation_block:

NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain A [imported] - Yersinia pestis (st
 C:Species: Yersinia pestis
 C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001
 C:Accession: AG0393
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M
 deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G
 11, M.; Rutherford, K.; Simmons, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell
 Nature 413, 523-527, 2001
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11586560
 A:Accession: AG0393
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-447 <KDR>
 A:Cross-references: GB:AL590842; PIDN:CAC92475.1; PID:g15981176; GSPDB:GN00175
 C:Genetics:
 A:Gene: ngrA
 C:Superfamily: Vibrio alginolyticus sodium-translocating NADH dehydrogenase (ubiquino
 C:Keywords: oxidoreductase

Quality: 1468.50 Length: 447
 Ratio: 3.814 Gaps: 1
 Percent Similarity: 86.130 Percent Identity: 62.416

Alignment block:

US-09-303-518D-125 x SS1015 ..

Align seg 1/1 to: SS1015 from: 1 to: 446

```

1 ATGATTTAAATCAAAAAAGCTCTTAACCTGAGCCATCGCGGCGAGACCGGA 50
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1 MetIethrIleIyLysGlyLeuAspLeuProIleAlaGlyThrProSe 17
51 GCAACCCGTTTACGACGGCCCGCCATTACGAAAGTCGCGTTGGCG 100
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
17 IclnValIleAsnAspGlyLysThrIleLysValAlaIleuLeuGlyG 34
101 AAGATATGCGGCTATGCGCCCTCGATGAAAGTCAGGAAGGCGATGCC 150
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
34 IuGlyIyValGlyMetIarProThrMetHisValArgValGlyAspLu 50
151 GTCAAAAAGGCCAAGTGTCTTGAAGACAAAAGAACATCCGGCGTGGT 200
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
51 ValLysLysAlaGlnValLeuPheGlnAspLysLysAsnProGlyVal 67
201 GTTACTCGCGCGCTTCAGCAAAATCGCGCATTCACCGTGGCGAA 250
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
67 sPheThrIarProIleAlaGlyLysValIleGluValAsnArgLysAla 84
251 AGCGGCTCTCAGTCAGTGTGATGCCCTGGAAGCAACGACGAGCAATC 300
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
84 ysaIyValLeuGlnSerValValIleGluValAlaGlyGluGlnVal 100
301 GAGTTTGAACGCTACGACCTGAGCGCGTGCACAACTTAAGCGCGAGA 350
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
101 ThrPheAspLysPheGlnAlaIleAlaGlnLeuSerLysLeuAsnArgLys 117
351 AGTGGCGCGCAACCTGATCCAAATCGGTTTGGACTGCGCGTGGACCC 400
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
117 ILeuLysThrGlnLeuValAspSerGlyLeuThrAlaLeuArgThra 134
401 GTCCGTTACAGAAATTCCTGCGCTCGATGCGCGCGCTTGGCATCTTC 450
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
134 rgrProPheSerLysValProAlaIleGlnSerThrLysAlaIlePhe 150
451 GTCAATCGATGACACCAATCCGCTGGTGCACGCTTACGGTCAATTAT 500
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
151 ValThrIleMetAspThrAsnProIleuAlaIalysProGluLeuIleI 167
501 CAAGAAGCGCGGAGATTTCAACCGCGCTGTTGTTAGCCGTT 550
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
167 eaSnGluGlnGlnIleAlaPheIleAlaGlyLeuAspIleLeuSerAlaI 184
551 TGACGGAACGCAAAATTCATGTTTGAAGCAGCTGCGGAGACGTCGCG 600
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
184 eurThrGlnGlyLysValIyValCysLys...SerGlyThrSerLeuPro 199
601 TCGAANAATCTGCGCAACATCGAACAACATGAATTCGGCGCGCGATCC 650
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
200 ArgSerSerGlnSerAsnValGluGlnHisValPheAspGlyProHisP 216
651 TGCGGTTTGAAGCGACGACATTCATTCATTCAGCGCGTGGCGGCA 700
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
216 oAlaGlyLeuAlaGlyThrHisMetHisPheLeuIyProValAsnAlaG 233
701 ATAAACCGTGTGACCATCAATTAATCAAGATTAATTAATTCAGCGCGT 750
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
233 IuAsnValAlaIarPserIleAsnIyGlnAspValIleAlaPheGlyLys 249
751 TTGTTTGAACAGCGCGCTGGAACACCGAGCGGTGATTCGCTTAGTG 800
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
250 LeuPheLeuThrGlyLeuLeuIyThrAspArgValValSerLeuAlaG 266

```

```

801 TTCTCAAGTCAACAACCGCGCTCTTGCTACGCTTTGGTGGCGGAAG 850
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
266 yProValValAsnAsnProArgLeuValArgThrValIleGlyAlaSerL 283
851 TATCGCAAAATTTACTGCGGGAATGTTGATACACAGCAACCGCTGAT 900
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
283 euAspAspLeuThrAspAsnGlnLeuMetProGlyLeuValArgValIle 299
901 TCCGGTTCCGTTATGACGCGCGGATTAACAAGCGCGCCGATTAATT 950
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
300 SerGlySerValLeuThrGlyThrHisAlaThrGlyProHisAlaIyI 316
951 GCGACCTTACACATCAATTCCTTATCGAAGAACGCGGACGGAAG 1000
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
316 uGlyArgIyThrHisGlnGlnValSerValLeuArgGlyArgGlyLysG 333
1001 AGCTGTTGCGCTGGGTTGCGCGCACCGGACAAATATCTCATCACGCT 1050
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
333 IuLeuPheGlyThrPalMetProGlyLysAsnLysPheSerValThrArg 349
1051 ACAACCCCTGGCCATTTCTGAAAAACAACCTTCAAGTTCAACACAGC 1100
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
350 SerPheLeuGlyHisValPheLysGlyGlnLeuPheAsnMetThrTh 366
1101 CGTCAACGGCGGACCGCGCCATGTCGATGTCGATTCAGACGCGG 1150
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
366 rThrAsnGlySerAspArgSerMetValProIleGlyAsnIyGlnValG 383
1151 TGATGCGCTTGGAATATCTGCGCCACCTGCTTTGCGCGATTAATCTGC 1200
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
383 alMetProLeuAspMetGluProThrLeuLeuArgAspLeuGlyAla 399
1201 GCGCATACGACAGACGCGGACGATTTGGTGTCTGGAATGACAGAGA 1250
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
400 GlyAspThrAspSerAlaIleAlaLeuGlyAlaLeuGlnLeuAspGluG 416
1251 AGACCTGCGTTTGTGACGCTTCTGCGCGGCGCAAAATACAGCGCC 1300
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
416 uAspLeuAlaLeuGlyThrPheValCysProGlyLysIyGlyIyGlyT 433
1301 CGCTGTGCGCAAACTGCTGCAAAACATTAAGAGGAAGCC 1341
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
433 hrLeuLeuArgGlnLysLeuAspThrIleGlnLysGlnGly 446

```

seq_name: p1r2:G82094

seq_documentation_block:

sodium-translocating NADH dehydrogenase (ubiquinone) (EC 1.6.5.-) nqrA VC2295 [import
 C:Species: Vibrio cholerae
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 16-Feb-2001

C:Accession: G82094

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.
 chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000
 A:Title: DNA Sequence of Both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833

A:Accession: G82094

A:Molecule type: DNA

A:Residues: 1462 <HEI>

A:Cross-references: GB:AE004300; GB:AE003852; NID:9656850; PIDN:AAF95439.1; GSPDB:GN

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Map position: 1

C:Superfamily: Vibrio alginolyticus sodium-translocating NADH dehydrogenase (ubiquino

C:Keywords: oxidoreductase

alignment_scores:

Quality: 1455.50 Length: 447
 Ratio: 3.810 Gaps: 1

Percent Similarity: 85.459 Percent Identity: 61.074

Alignment_block:

US-09-303-518D-125 x G82094

Align seg 1/1 to: G82094 from: 1 to: 462

```

1 ATGATTAATAAAGGCTTAACCTCCATGCGGCGACGACCGA 50
  |||||
17 MetIleThrIleLysGlyLeuAspLeuProIleAlaGlyThrProSe 33
51 GCAAGCCGTTAGCAGCGCGCCATTACCGAAGCGCTTCTGGCG 100
  |||||
33 GlnIleLeuSerAspGlyLysAlaIleLysLysValAlaLeuLeuGly 50
101 AAGATATGCGGTATGCGCCCTCGATGAAGTCAAGGAAGCGATGCC 150
  |||||
50 IuGIuTyValGlyMetArgProThrMetHisValArgValGlyAspGlu 66
151 GTCAAAAAAGCCAGTCTGTTGAAGACAAAAAGAAATCCGGCGCT 200
  |||||
67 ValLysLysAlaGlnIleLeuPheGluAspLysAsnProGlyValLys 83
201 GTTTCGCGCGCGCTTACGCAAAATGCGCGCGATTCACCGTGGCGAAA 250
  |||||
83 spherinSerProValSerGlyLysValValGlnIleAsnArgGlyAla 100
251 AGCGGCTACTTCAGTCACTGCTGATGCGCTTGAAGCAACGACGAAATC 300
  |||||
100 yArgValLeuGlnSerValValIleGluValAlaGlyAspAspGlnVal 116
301 GAGTTTGAACGCTACGCACTGAAGCGCGTGGCAAACTTAACGCGCGAGA 350
  |||||
117 ThrPheAspLysPheGluAlaAsnGlnLeuAlaSerLeuAsnAlaGlyAspAl 133
351 AGTGGCGCCACACTGATGCAATCCGCTTGTGACTGCGCTGGCGCAACC 400
  |||||
133 alIeLysThrGlnLeuValGlnSerGlyLeuThrPheAlaPheArgThrA 150
401 GTCCGCTGACAAATTCCTCGCGTGCATGCCAGCCGTTGCCATCTTC 450
  |||||
150 rGProPheSerLysValProAlaIleAspSerThrSerGlnAlaIlePhe 166
451 GTCAATGCGATGACACCAATCCGCTGGCTGGCGACCTAGGTCATAT 500
  |||||
167 ValThrAlaMetAspThrAsnProLeuAlaAlaGluProIleValAla 183
501 CAAAGAAGCGCGCGAGATTCAAACGCGCGCTTGGTATTTGAAGCCGT 550
  |||||
183 eaGlnGlnGlnSerGlnAlaPheValAlaGlyLeuAspValLeuSerAla 200
551 TGACGCAAGCGCAAAATCATGTTTGTGAAGCAGCTGGCGCGACAGCTGCCG 600
  |||||
200 eutThrGlyLysValTyValTyValCysLys...LysGlyThrSerLeuPro 215
601 TCTGAATAAGTCTCCACATCGAACAACATGTAATTCGGGCGCGCCATCC 650
  |||||
216 ArgSerGlnGlnProAsnValGlnGlnHisValPheAspGlyProHisPr 232
651 TGCCGCTTGTAGTGCACGACATTCATTTTCATCGACGCGGTGGCGGGA 700
  |||||
232 oAlaGlyLeuAlaGlyThrHisMetHisPheLeuTyProValSerAlaA 249
701 ATAAACCGTGTGGACCATCATTTATCAAGATGTAATTCACCTTGGCCGT 750
  |||||
249 sPhaIValAlaIlePheSerIleAsnTyGlnAspValIleAlaValGlyGln 265
751 TTGTTTCAACAGCGCGCTGTGAACACGAGCGCGTATGCGCTAGTGG 800
  |||||
266 LeuPheLeuThrGlyGlnLeuTyThrGlnArgValValSerLeuAlaGln 282
801 TTCTCAAGTCACAAACCGCGCTTGTGCGTACCGTTTGGTGGCAAG 850
  |||||

```

```

282 yProValValAsnLysProArgLeuValArgThrValMetGlyAlaSerL 299
851 TATCGCAAAATTACTCGCGGCGATTTGGTTGACACAGACAAACCGCTGAT 900
  |||||
299 euGIuGlnLeuValAspSerGlnIleMetCProGlyGlnValArgIleLe 315
901 TCCGTTTCGGTATTTGAACGCGCGGATTTACACAAGCGCGCGCATTTATT 950
  |||||
316 SerGlySerValLeuSerGlyThrLysAlaThrGlyProHisAlaTyLe 332
951 GGAAGCTTACCAACATCATGATTTCCGTTATTCAGAAGAAAGCCGAGCAAG 1000
  |||||
332 uGlyArgTyHisLeuGlnValSerValLeuAspGlnGlyArgAspLysG 349
1001 AGCTGTTCGGCTGGTTCGCGCGCGACGCGCAAAATACATCATCACGCGT 1050
  |||||
349 IuLeuPheGlyTyPAlaIleMetProGlyLysAsnLysPheSerValThrArg 365
1051 ACAACCCCTCGGCCATTTCTGAAAACAAACCTTCAGATTCAACACACG 1100
  |||||
366 SerPheLeuGlnHisLeuPheLysGlyGlnValTyLysMetThrThrTh 382
1101 CGTCAACGCGCGCGCGCGCGCATGCTGCGATTTGCTTACGACGCGG 1150
  |||||
382 rThrAsnGlySerAspArgSerMetValProIleGlyAsnTyGlnLysV 399
1151 TGATGCCCTTGATATCTCTGCCACCTGCTTGTGCGCGATTAATTCGTC 1200
  |||||
399 alMetProLeuAspMetClnProThrLeuLeuAspLeuAspLysAla 415
1201 GCGGATACCGACAGCGCGCGCATTTGGGTTGCTTGAATTTGACGAAGA 1250
  |||||
416 GlyAspSerAspSerAlaValAlaArgLeuGlyAlaLeuGlnLeuAspGln 432
1251 AGACCTCGCTTTGTCAGCTTCGCTGCGCGCGGCAAAATACGAATACGCGC 1300
  |||||
432 uAspLeuAlaLeuCysTrnPheValCysProGlyLysTyGlnTyGlnGly 449
1301 CGCTGTTCGCAAAAGTCTGGAACCAATTCGAAGGAAGGC 1341
  |||||
449 IuLeuLeuArgGlnLysLeuAspLysIleGlnLysGlnGly 462

```

seq_name: pir2:H83272

seq_documentation_block:

sodium-translocating NADH dehydrogenase (ubiquinone) (EC 1.6.5.-) nqr1 chain PA2999 [C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 16-Feb-2001
 C:Accession: H83272
 R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Adam, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
 A:Reference number: A82950; M01D:2043737
 A:Accession: H83272
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-445 <STO>
 A:Cross-references: GB:AE004724; GB:AE004091; NID:99949083; PIDN:AA06387.1; GSPDB:GN000001
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: nqrA; PA2999
 C:Superfamily: Vibrio alginolyticus sodium-translocating NADH dehydrogenase (ubiquinone)
 C:Keywords: oxidoreductase

alignment_scores:

Quality:	1353.00	Length:	447
Ratio:	3.627	Gaps:	1
Percent Similarity:	83.445	Percent Identity:	58.613

alignment_block:

US-09-303-518D-125 X H83272 .

Align seg 1/1 to: H83272 from: 1 to: 445

```

1  ATGTAATAATCAAAAAAGGCTTAACCTGCGCATCGCGGACAGCCGA  50
1  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1  MetLeuLysIleLysAsnArgLysLeuAspProIleSerGlyAlaProG1  17
51  GCAAGCGGTTTACACGCGCCGCGCATTTACCAGAAAGTGCGTGGTGGG  100
17  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
17  uGlnArgIleGlnAlaIleAlaArgProValArgSerValAlaLeuLeuLeu  34
101  AAGAATATGGCCGGTATGCGCCCTCGATGAAGTCAAGAGGCGATGCC  150
34  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
34  heAspTyrHisGlyMetLysProThrMetAlaValGlnValGlyAsnArg  50
151  GTCAAAAAAGCCCAAGCGCTGTTTGAAGACAAAAAGATCCGGCGCGCT  200
51  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
51  ValLysLeuLeuGlnValLeuAlaLeuPheThrAspLysLysAsnProSerValSe  67
201  GTTTACTGCGCGCGCTTACAGCAAAATCGCCGCGATTCACCTGCGGAAA  250
67  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
67  rTyrThrAlaProG1ValaGlyValValaSerAlaIleHisArgGlyGln  84
251  AGCGCGACTGTAGTCAAGTGTGATGCGCGTGTAAAGCAACAGCAAAATC  300
84  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
84  ysAspTyrAlaLeuGlnSerValValaIleAspLeuAsnSpLyspGlnGlnLeu  100
301  GAGTTTGAACGCTACGCACCTCAAGCGCGTCAAACTTAAGCGGCGCAAGA  350
101  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
101  GluPheAlaArgTyrProAlaAspLysLeuAlaThrLeuSerAlaLeu  117
351  AGTGGCGCCCAACCTGATCAATCCGCGTTGTGGACTGCGTGCAGCC  400
117  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
117  nValaArgAspAsnLeuLeuGlnSerLysLeuTyrThrAlaLeuArgTbra  134
401  GTCCGCTTACGAAATTCCTGCGCGTGCATGCGCGAGCGCGTTCGCATCTTC  450
134  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
134  rgProPheSerLysValProAspProGlnSerSerProSerSerIlePhe  150
451  GTCAATGGCATGACACCAATCCGCGTGGCGCGCACTTACGGTCAATAT  500
151  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
151  ValThrAlaIleAspThrGlnProLeuAlaIleAspProGlnValValaI  167
501  CAAAGAAGCGCGGAGGATTTCAACGCGCGCGCTGTGGATTTGAGACCGTT  550
167  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
167  eAlaGlnGlnGlnGlyAlaPheGlnAlaGlyLeuThrValLeuGlnGly  184
551  TGACCGAAACGCAAAATCCATGTTTGTAAAGCAAGTGGCGAGACGTGGCG  600
184  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
184  euAla.....ArgValPheLeuGlyLysAlaGlnGlyValSerLeuPro  198
601  TCTGAATAATGCTGCCACATGTGAACACATGATATGGCGCGCCGCATCC  650
198  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
198  GlyGlnAlaLeuSerGlyValaThrAlaGlnAlaPheSerGlyProHisPr  215
651  TGCCGTTTGAATGACGACGACATTCATTCATTCACGACGCGGTGGCGGGA  700
215  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
215  oAlaGlyLeuProGlyThrHisIleHisPheLeuAspProValaGlyAlaG  232
701  ATAAACCGGTGGACCATCATTAATCAAGATGTAAATTTACATTTGGCGGT  750
232  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
232  LysLysSerValaTyrPheAsnLeuAsnTyrGlnAspValaIleAlaIleGlyLys  248
751  TTGTTTGAACAGCGCGCTGTCAACACGAGCGCGCTGATTCGCCATAGTGG  800
248  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
248  LeuPheThrThrArgLysGlnLeuThrPheGlnArgValaIleAlaLeuAlaG  265
801  TTTCGAAGTCAACAAACCGCGCGCTTTCGCGATTCGTTTGGGTGGCAAG  850
265  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
265  yProValaValaGlnLysProAlaArgLeuValaArgThrArgLeuGlnGlyAlaAsnL  282
851  TATCGCAAAATTACTCGCGGCGGAATGTGTTTACACAGACAGCAACCGCGTATT  900

```

```

282      eaaergshleuAlaIaGlyGlnLeuGlnProGlyAsnSnrArgLeuTle 298
901  TCGGGTTCCGGTATTGAAAGCGCGGATTTACAAAGCGCGCGCAATTAATT 950
299  SetGlySerValLeuLeuGlyGlnArgThrAlaHisGlyAlaTyrAlaTyrLe 315
951  GGAACGGTATACACAAATACAGATTTGGTATTCGAAGAAAGCGCGCACCAAG 1000
315  uGlyArgTyrHisLeuGlnLeuSerCysLeuGlySerGlyAspGlnArgG 332
1001  AGCTGTTCGGCTGGGTTCGCGCGCAAGCGCGCAAAATACATCCATCAGCGGT 1050
332  lPheLeuHisTyrLeuAlaGlyAlaGlyValGlnIleGlyHisSerLeuLeuAsn 348
1051  ACAACCGTGGCGCATTCCTGTGAAAAACAACATCTTCAAGTTTCACACAGC 1100
349  ValPheValSerArgLeuLeuLeuGlyGlyAspArgPheAlaPheThrThrSe 365
1101  CGTCAACGGCGGCGAGCAAGCGCGCATGGTGGCCGATTTGGTACTTACAGCGCG 1150
365  rThyAsnGlySerProArgAlaMetValProValGlyAsnTyrGlnAlaVal 382
1151  TGAATGCCCTGGATATCTGCGCCACCCCTGGCTTTCGCGGATTAATGATGC 1200
382  AlMetProLeuAspIleLeuProThrGlnLeuAlaTyrLeuIleVal 398
1201  GCGGATACCGAAGCGCGCAAGGATTCGGGTGGTGGATTCGACGCAAGA 1250
399  GlyAspThrGlnMetAlaGlnIlyLeuGlyAlaLeuGlnLeuAspGlnI 415
1251  AGACCTCGCTTTGTGACAGCTTCGTTGCTGCCCGCGCAAAATACGAATACGGCC 1300
415  uAspLeuAlaLeuCysSerTyrValCysAlaGlyLeuTyrGlnIleTyrGly 432
1301  GCGTGTGGCGAAGATCGTGGAAACCAATTCGAAGGAAGCG 1341
432  rTlleLeuAlaTyrAspAsnIleuAlaArgIleGlnIleGlnIleGly 445

```

```

seq_name: plir2:C81751
seq_documentation_block:
  Probable sodium-translocating NADH dehydrogenase (ubiquinone) (EC 1.6.5.-) alpha chain
C:Species: Chlamydia muridarum, Chlamydia trachomatis Mopn
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 16-Feb-2001
C:Accession: C81751
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hickey,
  C.C.; Dodson, R.; Gwinn, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzberg,
  A.; et al.: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39
A:Reference number: A81500; MUID:20150255
A:Accession: C81751
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-465 <TE>
A:Cross-references: GB:AE002269; GB:AE002160; NID:g7190041; PIDN:AAF38896.1; PID:g719
A:Experimental source: strain NigG (Mopn)
C:Genetics:
A:Gene: TC0002
C:Superfamily: Vibrio alginolyticus sodium-translocating NADH dehydrogenase (ubiquinone)
C:Keywords: oxidoreductase

```

```
alignment_scores:
  Quality: 471.00
  Ratio: 1.581
  Percent Similarity: 64.224
  Length: 466
  Gaps: 14
  Percent Identity: 29.095
```

```
alignment_block:
US-09-303-518D-125 x C81751
```

Align seg 1/1 to: C81751 from: 1 to: 465


```

3  lIeThVaLaSnArGdLyLeuAsPLeuSerLeuGInGlySerProLySGl 19
51  GCAAGCCGTTTACGAC.....GGCCCGCCCATTCACGAGTCGGTTCG 94
19  uSerGlyPheThyAsnLySIlLeAsPProGluPheValSerIleAsPLeu 35
95  TTGGCGAAGATATGCGCGTATGCGCCCGCCCGATGCAAGTCAGAGAGC 144
36  ....ArgProPheGlnProLeuSerLeuLySLeuLyValGluGlnGly 50
145  GATGCCGTCAAAAAGGCCAGTGTGTTTGAAGCAAAAGATCCGGG 194
51  AsPaLaValLySSeRgLyAlaProIleAlaGluTyLyShISpheProAs 67
195  CGTGSGTTTACTGCGCGCGCTTACGCAAAATGCGCGCGATTCACCGTG 244
67  nThyTylIleThySerHisValSerGlyValAlaThraIleAlaArgArG 84
245  GCGAAAGCGCGCTTACTGACGTACGTGATG.....GCCGTGAAGGCAAC 291
84  LySnLyAsPLeuSerLeuAsPValIleIleLySlyThProGlyPro 100
292  GACGAAATCGAGTTTGAACGCTACGACACCTGACGCGTGGCAACTAG 341
101  ThSerThrGluTy.....ThTyAsPLeuGlnThyLeuSe 113
342  CGCGAAGACGTGCGCGCGCAACTGATTCATTCGCTTGTGACTGCGC 391
113  rArgSerAsPLeuSerGlnIlePheLySlyAsnGlyLeuPheAlaLeuI 130
392  TGCCGACCGCGTGGTTCAGCAAAATTCCTGCGCGTGCATGCCGAG...CCG 438
130  leTySglnAlaProPheAsP....IleProAlaIleProIlnGlnThPro 145
439  TTGCGCATTTCTGTCATGCGATGACACCAATCCGCTGCGCGCGCGCC 488
146  ArgSArValPheIleAsnLeuAlaAsPAsnArgProPheThrProSerPr 162
489  TACGCTATATC.....AAGAAGCCGCGAGGATP 520
162  oGluLySnIleLeuAlaLeuPheSerSerArgGlnGlnGlyValr 179
521  TCAAAACCGCGCTGTTGATGAGCGGTGAGCCGCAAGCAAAATCCAT 570
179  heValValGlyValArgAlaIleAlaLySLeuPheGlyLeuArgProHis 195
571  GTTGTGAAGCGACGTGCGCGAGACGTGCGCTGAAATGCTGCCAATC 620
196  lIeValrPheArgSArValrGlyLeuThrLeuProThrGlnGlnLeuLySthrI 212
621  C...GAACACATGATGCGCGCGCGCGCGCATCCTGCGCGTGGATGGCA 667
212  eAlaHisLeuHisThrValSerGlyProPheProSerGlySerProSerI 229
668  CGACATTCATTCATTCAGACCGCGTGGCGCGCGCATATA...ACCGTGG 714
229  leHisIleHisSerValAlaProIleThraSngLySgluValAlaPhe 245
715  ACATGATATTCATGACATGATTCATTCATTCGCGCGCTTGTGCAACAG 764
246  ThrLeuSerPheGlnAsPValIleuThrIleGlnHisLeuPheLeuLySgl 262
765  CGGTGCAACACCGAGCGCGGTGATGCGCGTGGATGGTTCGCAAGTACA 814
262  yArgIleLeuHisGlnGlnValThraIleLeuAlaGlyThraIleAlaLyS 279
815  AACCG.....CGCGTTCGCTACCGCTTGTGGTGGCAAGATATCGCA 858
279  eSerLeuAlaArgTyValAlleThrThrLySglYAlaSerPheSer 295
859  ...ATTACTGCGCGCGAATGGTTCACACAGCAACACCGCTGATTCGCG 905
296  leuIleAsnLeuAsnAsPleSerAsPAsnSprh...LeuIleSerGln 311

```

```

906  TTGGTATTCAGCGCGCATTCACAGAGCGCCGACGAT...TATTGG 952
311  yAsPProLeuThrGlyArgLeuCySlySglnGlnGluProPheLeuG 328
953  GACGCTACCCATTCAGATTTCCGTTATCGAAGAGCGCGCAAGAG 1002
328  LyPheArgAsPHisSerIleSerValLeuHisAsnProThrLySArgGln 344
1003  CTGTTGCGTGGGTTGCGCGCGACCGCAAAATATCCATTCACGCGTAC 1052
345  LeuPheSerPheLeuAlaGlyIleGlyPheAsnLySProIlnPheThrLyS 361
1053  AACCTCGCGCATTTCTGTAATAAACAATCTTCAGATTCAC.....A 1096
361  rTyTyLeuSerGlyPhePheLySlySArgThrTyThraSProAsPrt 378
1097  CAGCGCTCAACGCGCGACGCGCGCATGTCGCGCATGTGATCTACAGAG 1146
378  hraSLeuHisGlyGluThraArgProIleIleAsPThraSPlleTyAsP 394
1147  CGCGTATCCCTTGGATATCCTGCGCACCCCTGTTTGGCGGATTTAT 1196
395  LySValMetProMetArgIleProValValProLeuIleLySAlaValI 411
1197  CGTGGCGCATTACCGACAGCGCGCGACGATGGGTTGCTTGAATTCGACG 1246
411  eThrLySAsnPheAsPLeuAlaAsnGlnLeuGlyPheLeuGlnValLySg 428
1247  AAGAAGACCTCGCTTGTGACCTTCGTCGTCGCGCGCAATACAGATAC 1296
428  LyGlnAsPheAlaLeuProThrIleLeuAsPProSerIleThyGlnMet 444
1297  GCGCGCTGTTGCGCAAGTCTGGAACCATTCGAAGAGAA 1338
445  leuThrIleValLySgluSerLeuIleGlyTyAlaLySgln 458

```

seq_name: p1r2:E72040

```

seq_documentation_block:
  probable sodium-translocating NADH dehydrogenase (ubiquinone) (EC 1.6.5.-) alpha chain
  C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
  C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 16-Feb-2001
  C:Accession: E72040; G81623
  R:Kalmann, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood,
  Nature Genet 21: 385-389, 1999
  A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
  A:Reference number: A72000; M01D:99206606
  A:Accession: E72040
  A:Molecule type: DNA
  A:Residues: 1-467 <RNA>
  A:Cross-references: GB:AE001656; GB:AE001363; NID:g4377047; PIDN:AD18882.1; PID:g437
  A:Experimental source: Strain GW1029
  R:Read, T.D.; Brinham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke
  C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe
  Nucleic Acids Res. 28, 1397-1406, 2000
  A:Title: Genome sequences of Chlamydia trachomatis Morn and Chlamydia pneumoniae AK39
  A:Reference number: AB1500; M01D:20150255
  A:Accession: G81623
  A:Molecule type: DNA
  A:Residues: 1-467 <RNA>
  A:Cross-references: GB:AE002164; GB:AE002161; NID:g7188939; PIDN:AAF37899.1; PID:g718
  A:Experimental source: strain AR39, HL cells
  C:Genetics:
  A:Gene: nqrA; CP00002
  C:Superfamily: Vibrio alginolyticus sodium-translocating NADH dehydrogenase (ubiquino
  C:Keywords: oxidoreductase

```

```

alignment_scores:
  Quality: 464.50      Length: 464
  Ratio: 1.585        Gaps: 15
  Percent Similarity: 63.147  Percent Identity: 30.172

```

alignment_block:
US-09-303-518D-125 x E72040 ..

Align seg 1/1 to: E72040 from: 1 to: 467

```

4 ATTAATCAAAAAAGCTTAACCTCCCATCGCGGCGAGACCG...GA 50
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
3 IleThrValAsnArgGlyLeuAspLeuSerLeuGlnGlySerProLys 19
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
51 GCAACCCCTTACGAC.....GCCCGGCCATTACCGAAGTCGCGCTGC 94
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
19 uSerGlyPheThrAsnLysIleAspProGluPheValSerIleAspLeu 35
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
95 TTGGCGCAAAATATGCGCGGATGCGCCCTCGATGAAGTCAAGAGAGC 144
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
36 ....ArgProPheGlnProLeuSerLeuLysValGluGlnGly 50
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
145 GATGCCGCTCAAAAAAGCCAAAGTCTGTTTGAACAACAAAAGATCCGGG 194
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
51 AspAlaValAlcysSerGlyAlaProIleAlaGluThrLysHisPheProAs 67
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
195 GGTGGGTATTACTGCGCGGCTGAGCAAAATGCGCGCATTCACCGTGC 244
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
67 mThrTyrlIleHisSerHisValSerGlyValAlaThrAlaIleArgArg 84
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
245 GCGAAAAAGCGCGTACTTCAAGTCAAGTCTGAT...GCCGTGAAGGCAAC 291
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
84 LysHisLysArgSerLeuLeuAspValIleIleLysLysThrProGlyPro 100
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
292 GACGAATAGAGTTTGAAACGCTACCCACTGAGAGCGCTGGCAACTTAA 341
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
101 ThrSerThrGluTyr.....ThrTyrlAspLeuGlnThrLeu 113
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
342 CGCGAAGAGAGTGGCGCGCAACCTGATCCAAATCCGTTTGATGATGCGC 391
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
113 ArgSerArgLeuSerGluIlePheLysGluAsnGlyLeuPheAlaLeu 130
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
392 TGGCAACCGCGCTGTTGATGAGCAAAATCTCGCGCTGATGCGGAG...CCG 438
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
130 LeuLysGlnArgProPheAsp...IleProAlaIleProThrGlnThrPro 145
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
439 TTGGCATATGTCATGATGATGCAACCAATCCGCTGGCGGCGGACCC 488
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
146 ArgAspValAlaPheIleAsnLeuAlaAspAsnArgProPheThrProSer 162
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
489 TACGGTCATTATC.....AAGAGCGCGCGAGGATT 520
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
162 GGUlyuHisLeuAlaLeuPheSerSerArgGluGlnGlyPheTyrVal 179
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
521 TCAACCGCGCGCTGTTGATGAGCAAAATCTCGCGCTGATGCGGATTCAT 570
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
179 heValValGlyValArgAlaIleAlaLysLeuPheGlyLeuArgProHis 195
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
571 GTTGTGAAGCAGCTGGCGGAGAGTGGCTGAAATGCTGCCAAT 620
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
212 eAlaHisLeuHisThrValSerGlyProPheProSerGlySerProSer 229
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
668 CGGCAATTCATTCATGAGCGCGTGGCGGCGCAATTA...ACCGTGG 714
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
229 LeuHisIleHisSerValAlaProIleThrAsnGluLysGluValAlaPhe 245
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
715 ACATCAATTAACAAGATTAATTAACATGCGCGTGTGTTGCAACAG 764
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
246 ThrLeuSerPheGlnAspValLeuThrIleGlyHisLeuPheLeuLysG 262
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
765 CCGTGTGAACAGCGCGCTGATGCGCTGAGGTGCTGCAAGTCAAG 814
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
262 yArgIleLeuHisGlnGlnValThrAlaLeuAlaGlyThrAlaLeuLys 279

```

```

815 AACCG.....CGCTCTGCTACCGTTTGGTGGCAAGATTCGCA 858
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
279 eSerLeuArgArgTyrValIleThrLysGlyAlaSerPheSer 295
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
859 ...ATTACTGGGCGCAATGTTGATGACACAGACACCGCGTGAATTCGG 905
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
296 LeuLeuAsnLeuAsnAspIleSerAspAsnAspThr...LeuIleSerG 311
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
906 TTGCGTATTGACGCGCGCATTCACAGCGCGCAGCAT...TATTGG 952
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
311 yAspProLeuThrGlyArgLeuLysLysGluGluGluProPheLeuG 328
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
953 GACGCTACCAACAATCAGATTCGTTTCGATTCGAAGAGCGCGCAAGAA 1002
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
328 lPheAlaArgAspHisSerIleSerValLeuHisAsnProThrLysArg 344
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
1003 CTGTTGCGGCTGTTGCGCGCGACCGCAAAATCTCATCAACGCGTAC 1052
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
345 LeuPheSerPheLeuArgIleGlyPheAsnLysProThrPheThrLys 361
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
1053 AACCTCGCGCATTTCTGAAAAACAACCTTCAATTCAC.....A 1096
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
361 rTyrlLeuSerGlyPhePheLysLysArgThrTyrlThrAsnProAsp 378
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
1097 GAGCGCTACAGCGCGCGACCGCCATGATGCGCATTTGATTCACAG 1146
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
378 hAsnLeuHisGlyGluThrArgProIleIleAspThrAspIleTyrAsp 394
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
1147 CGCGTATGCGCTTGATATCTGCGCCACCGCTGTTTGGCGCATTAAT 1196
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
395 LysValMetProMetArgIleProValAlaProLeuLeuLysAlaVal 411
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
1197 CGTGGCGCATACCGACAGCGCGGCGCATGTTGCTGGAATTTGAGAC 1246
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
411 eThrLysAsnPheAspLeuAlaAsnGluLeuGlyPheLeuGluValLysG 428
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
1247 AAGAGACCTCGCTTGTGACGCTTCGTCGCGCGGCAATTAACATAC 1296
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
428 lYglAspPheAlaLeuProThrLeuIleAspProSerLysThrGluMet 444
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
1297 GCGCGCTGTTGCGCAACGCTGCAAAACATTTGAGAGAGAA 1338
  ||| :||| :||| :||| :||| :||| :||| :||| :|||
445 LeuThrIleValLysGluSerLeuIleGluTyrAlaLysGlu 458
  ||| :||| :||| :||| :||| :||| :||| :||| :|||

```

seq_name: p1r2:F71489

seq_documentation_block:

sodium-translocating NADH dehydrogenase (ubiquinone) (EC 1.6.5.-) alpha chain [siml1a
C:Species: Chlamydia trachomatis
C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 02-Mar-2001
C:Accession: F71489
R:Stephens, R.S.; Kalman, S.;ammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitche
Science 282, 754-759, 1998
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia t
A:Reference number: A71570; M01D:99000805
A:Accession: F71489
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-465 <ARN>
A:Cross-references: GB:AE001334; GB:AE001273; NID:93329078; PIDN:AMC68238.1; PID:9332
A:Experimental source: serotype D, strain UW-3/Cx
C:Genetics:
A:Gene: nqrA
C:Superfamily: Vibrio alginolyticus sodium-translocating NADH dehydrogenase (ubiquino
C:Keywords: oxidoreductase

alignment_scores:

quality:	446.00	length:	464
ratio:	1.517	gaps:	14
percent similarity:	63.362	percent identity:	28.233

alignment block:

US-09-303-518d-125 x F71489 ..

Align seg 1/1 to: F71489 from: 1 to: 465

```

4 ATTTAAATCAAAAAAGGTCTAATACCTGACCATCGGCGGCGAGACCGAGACA 53
  ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
3 IIEVALVATSerArgGlyLeuAspSerLeuLysGlyAlaProLysG61 19
54 AGCCCTTACGAGCGC.....CCGCGCATTTACCGAAGTCGCGTTGC 94
  : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
19 uSerGlyPheCysGlyLysValAspProThrTyrValSerValAspLeu. 35
95 TTGGGAGAAATATGCCGTATGCCCGCCCTCGATGAAGAAAGTAAGAGAGAGC 144
  : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
36 ....ArgProPheAlaProLeuProLeuGlyValLysValThrProLu 50
145 GATCCCGTCAAAAAGGCCAAGTGTGTTGAGACAAAGAAAGATCCGGG 194
  ||| ||| : : : : : ||| ||| ||| ||| ||| |||
51 AspGlnValThrAlaGlySerProLeuAlaGlyLysLeuPheSerG1 67
195 CGTGGCTTAACTGCGCGCGCTTACGCAAAATCGCGCATTCACCGTG 244
  ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
67 yAlaPheLeuThrSerProValAspGlyGluValGluLeuArgArg 84
245 GCGAAAGCGGCTACTTCAGTCACTCGTGTGATGCCGTTGAA...GCGAAC 291
  ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
84 LysnLysArgAlaLeuLeuGluValLysLysLysProGlyLle 100
292 GACCAATTCAGATTGGAACGCTACGACACGCAACGCGTGGCAAACTAAG 341
  : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
101 SerGlnThrLysPheSer.....TyrAspLeuGlnSerLeuTh 113
342 CGCGGAAGAGATGCGCGCAACGATGCAATCCGTTGTGACTGCGC 391
  ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
113 rGlnLysAspLeuLeuGluValPheLysLysGluLysLeuPheAlaLeuP 130
392 TGGCACCCTGCGTTCAGCAAAATTCCTCCGTCGATGCCGAG...CCG 438
  : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
130 heLysGlnAlaProPheAsp...IleProAlaLeuProThrGlnSerPro 145
439 TTGGCCATCTTCGATCGATGCGATGACACCAATCCGCTG...GCTGC 482
  : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
146 ArgAspValPheIleAsnLeuAlaAspAsnArgProPheThrProSerVa 162
483 CGACCCCTACGCTATTATCAAGAGACCGCGCAGAT.....T 520
  : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
162 IGIuLysHisLeuSerLeuPheSerSerLysGluAspGlyTyrTyrIleP 179
521 TCAAAGCGCGCTGTGTGATTTAGCGCGTTTGACGCAAGCAAAATCCAT 570
  ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
179 heValValGlyValGlnAlaIleAlaLysLeuPheGlyLeuLysProHis 195
571 GTTTGAAAGCAGCTGCGCGCAGACGCTGCGTGAATGTCGCAACAT 620
  : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
196 IleIleSerThrAspArgLeuThrLeuProThrGlnAspLeuValSerI1 212
621 C...GAACACATGATTCGCGCGCGCGCATCTCGCGCTTGAAGTGA 667
  ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
212 eAlaHisLeuHisThrIleAspGlyProPheProSerGlySerProSerT 229
668 CGCAGATTCATTTCATCGAGCGCGT...GGCGCAATTTAAACCGTGGG 714
  ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
229 hrHisIleHisIleAlaArgIleArgAsnGluArgAspValAlaPhe 245
715 ACCATCATTTATCAAGATGATTAATACCATTCGCGCTTTGTTGCAACAG 764
  ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
246 ThrIleSerPheGlnGluValLeuSerIleGlyHisLeuPheLeuLysG1 262
765 CGCTGTGAACACCGAGCGCGGTGATGCCCTAGGCTTCTCAAGTC.... 810
  ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
262 yheValLeuGlyGlnGlnIleValAlaLeuAlaGlySerAlaLeuProp 279

```

```

811 ..AACAAACCGCGCTCTTGCGTACCGTTTGGGTGCGAAAGATTCGCA 858
  : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
279 roSerGlnArgLysTyrLeuIleThrAlaLysGlyAlaSerPheSerAsp 295
859 ATTATGCGGGCGAATTTGTTGACACAGACAGAAC...CGCGTATTTCCG 905
  : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
296 LeuLeuProLysAspIlePheSerSerAspGluIleThrLeuIleSerI 312
906 TTGGGTATTAAACGCGCGCATTTACACAGAGCGCGCAGAT...TATTTGG 952
  ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
312 yAspProLeuThrGlyArgLeuLysLysGluGlnAsnProCysLeuG 329
953 GAGCGTACCAATCAGATTTCGTTATCGAAGAAAGCGCGCAGCAAGAG 1002
  ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
329 LysMetArgAspHisThrIleThrLeuLeuProAsnProLysThrArgGlu 345
1003 CTGTGGCTGGGTGGTGGCGCGACCGCAAAATCTCCATCAGCCGTTAC 1052
  ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
346 SerPheSerPheLeuArgLeuGlyTyrAsnLysLeuThrValThrArgTh 362
1053 AACCTCGCGCATTTCTGAAAAACAA.....CTCTCAAGTTCAACA 1096
  ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
362 rTyrLeuSerGlyPhePheLysArgLysArgValPheMetAspMetAspT 379
1097 CAGCGCTCAACGCGCGCGACCGCGCATGTGCGCATTTGTACTACAGAG 1146
  ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
379 hrAsnMetIleGlyGluLysArgProIleIleAspAlaGluIleTyrGlu 395
1147 CGCGTGATGCCCTTGGAATTCCTGCCACCCCTGCTTTGCGGATTTAT 1196
  ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
396 ArgValSerAlaIleProValProValAlaLeuIleIleLysAlaLeuG1 412
1197 CGTGGCGCATTCAGACAGCGCGCAGCATTTGCTTGGATTTGACAG 1246
  ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
412 uThrGlnAsnPheGluGluAlaCysArgLeuGlyLeuLeuValAlaP 429
1247 AAGAAGACCTGCTTTGTGACAGCTGCTGCGCGCGCAATTAAGAAATAC 1296
  ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
429 roGluAspPheAlaLeuProThrPheIleAspProSerLysThrGluMet 445
1297 GGCCCGCTGTGGCGCAAAAGTCTGGAACCATTTGAAAGGAA 1338
  : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
446 PheSerIleValLysGluSerLeuLeuArgTyrAlaLysGlu 459

```

seq_name: p1f2:E82252

seq_documentation_block:

```

RntC-related protein VC1015 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C/Species: Vibrio cholerae
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C/Accession: E82252
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.
  chardson, D.; Esmailova, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers
  1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
A/Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A/Reference number: A82035; MUID:20406833
A/Accession: E82252
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-801 <HEID>
A/Cross-references: GB:AE004183; GB:AE003852; MID:99655473; PIDN:AAF94176.1; GSPDB:GN
A/Experimental source: serogroup O1; strain N16961; biotype El Tor
C/Genetics:
A:Gene: VC1015
A:Map position: 1
C:Superfamily: unassigned ferredoxin 2[4Fe-4S]-related proteins; ferredoxin 2[4Fe-4S]

```

alignment_scores:

```

Quality: 158.00 Length: 472
Ratio: 0.721 Gaps: 16
Percent Similarity: 46.398 Percent Identity: 20.127

```

alignment_block:
us-09-303-518D-125 x E82252

Align seg 1/1 to: E82252 from: 1 to: 801

```

52 CAAGCCGTTTACGAGCGCCGCGCATTCAGCAATCGCTTGGTCGCGA 101
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
59 GlnProIleIleAsnAlaSerIleProAsnGlnLeuValLeuProLeu 75
102 AGAATATGCGCGTATGCGCCCTCGATGAAAGTCAGCAAGCGCATCCG 151
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
75 sglHisIleIleLysAlaGlyAspLeuValLysValGlyAspArg 92
152 TCAAAAAGGCGCAGTGTCTTTGAGACAAAAGAAATCCGGCGGTG 201
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
92 alleuLysGlyGlnProLeuThrGlnIleThrSerThrPheMetLeuPro 108
202 TTTACGCGCGCGCTTACGCAAAATCGCGCATTCACCGTGGCGAAA 251
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
109 IleHisAlaProThrSerGlyValIleSerAlaIle.....GluPr 122
252 CGCGGTCTCTCAG.....TCAGTGTGATG 277
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
122 oArgThrValAlaHisProSerGlyLeuSerGluLeuCysIleValLeu 139
278 CCGTTGAAGCAGCAGCAGAA.....ATGAGTTT 306
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
139 hrProAspGlnGlnGlnIleThrPheGlnLeuGlnProGlnProAspPhe 155
307 GAAGCTACGACGCTGAAGCGCTGGCAAACTTAAAGCGGCAAGAGTGC 356
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
156 GlnGlnLeuThrProGlnThrLeuLeuGlnLeuIleArgGlnAlaGly 172
357 CCGGAACCTGATCCATCCGCTTGTGACTGCGCTGCGACCCGTCCT 406
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
172 eSerGlyMetIleGlyAlaGlyPheProThrAla..... 183
407 TCAGCAAAATCTCGCGTGCATGCGCAGCGCTTCCGATCTTGTGAT 456
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
184 ..LysLysLeuGlnSerGlyLeuSerArgThrGlnIleLeuIleAsn 199
457 GCGATGACACCAATCCGCTGGTGCACGACCTTACGCTTATCAAGA 506
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
200 AlaAlaGlnCysGlnProGlyThrAlaAspAspValLeuMetArg 216
507 AGCGCGGAGGATTTCAAGCGGCGCTGTGATTCAGCGCTTGCACG 556
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
216 nTyrAlaHisGlnIleIleGlnGlyIleGlnIleValGlnHisIleLeu 233
557 AACGCAAAATCCATGTT..... 573
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
233 yspProLysLeuThrIleIleGlyIleGlnAspAsnLysProGlnAlaVal 249
574 .....TGTAAAGCAGCTGGCGCAGACGCTGCTCT..... 603
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
250 AlaAlaLeuGlnGlnAlaAlaGlnAspLysProMetValIleArgVal 266
604 .....GAAATATCTG 613
266 eProThrLysTyrProSerGlyGlyLysGlnLeuIleLysIleLeu 283
614 CCAACATCGAACAACATGATTCGCGCGCGCATCTCTGCGCTTGTGAGT 663
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
283 hrAsnLeuGlnValProLysGlyIleProAlaAspIleGlyLeu... 298
664 GGCACGCAATTCATTCATGAGCGCGTGGCGCGAATAAAACCGTGTG 713
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
299 .....MetValGlnAsnValGlySer..... 305
714 GACCATCATTAACAAGATTAATTACCATGCGCTTGTGTCACAG 763
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
306 .....LeuGlnAlaIleAlaArgAlaIleValHisG 316

```

```

764 GCGCTGTGAACACGAGCGCGTATTCGCCCTAGTGTGTTCAAGTCAAC 813
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
316 LysIleProLeuIleArgValAlaThrLeuThrGlyAspPheArg 332
814 AACCGCGCGCTCTTGGTACCGCTTGTGGTGGCAAGTATCGCAATTAC 863
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
333 LysProArgAsnValThrAlaLeuLeuGlyThrProValGlnAlaLeu 349
864 TCGCGCGCAATTGTTGACACAGACAC.....CGCGTATTCG 904
349 unsngIlePheGlyTyrLysAlaAspLysLysLeuProArgLeuIleMet 366
905 GTTCGCTATTAACGCGCGATTCACCAAGCGCGCATATTTTGGGA 954
366 LysIleProMetMetGlyPheThrLeuPro..... 375
955 CGCTACCAATCAGATTCGTTACGTAAGAAGCGCGCAGCAAGAGCT 1004
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
376 .....HisAlaGlnValProIle.....ThrArgThrAl 385
1005 GTTCGCGCTGGTGGCGCGCGCAGCGACAAATCTCCATCAGCTACAA 1054
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
385 aAsnCysIleLeuAlaPro.....ThrArgAsnG 395
1055 CCTCGCGCATTTCTGAAAAAACTCTTCAGTTCAACACACCGCTC 1104
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
395 LysLeu.....ThrSer 398
1105 AACGCGCGCGCGCGCGCATGTGTCGATGTTGACTTACGACGCGGTAT 1154
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
399 SerAspAsnGlnMetAlaCysIleArgCysGlyIncysAlaGlnAlaC 415
1155 GCGCTGTGATTCCTGCGCCACCTGCTTGTGGCGCATTTATCGTCGCG 1204
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
415 sProValSerLeuLeuProGlnGlnLeuGlnIleThrPheAlaLysAlaGln 432
1205 ATACCGACGACGCGCGCATGTTGGTGTGAAATTCGACGCAAGAAC 1254
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
432 LysPheAspLys.....CysGlnGlnLeuAsnLysPheLysAsp 443
1255 .....CTGCTTGTGACGCTTGTGTCGCCGCGGCAATTCAGAA 1295
444 CysIleGlnCysGlyAlaCysAlaThrValCysProSerGlnIle.... 458
1296 CGGCGCGCTGTGGCG 1311
459 ....ProLeuValGln 462

```

seq_name: p1r2:S39893

seq_documentation_block:

rnfC protein - Rhodobacter capsulatus
 C:Species: Rhodobacter capsulatus
 C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 17-Mar-2000
 C:Accession: S39893
 R:Schmehl, M.; Jahn, A.; Meyer zu Vilsendorf, A.; Hennecke, S.; Masepohl, B.; Schupp
 Mol. Gen. Genet. 241, 602-615, 1993
 A:Title: Identification of a new class of nitrogen fixation genes in Rhodobacter caps
 A:Reference number: S39892; M01D:94086454
 A:Accession: S39893
 A:Molecule type: DNA
 A:Residues: 1-519 <SCS>
 A:Cross-references: EMBL:X72888; NID:9435523; PIDN:CA51399.1; PID:9435525
 C:Gene: rnfC
 C:Superfamily: unassigned ferredoxin 2[4Fe-4S]-related proteins; ferredoxin 2[4Fe-4S]
 C:Keywords: iron-sulfur protein
 F:374-438/Domain: ferredoxin 2[4Fe-4S] homology <FBR>

alignment_scores: Quality: 150.00 Length: 454

Ratio: 0.676 Gaps: 18
Percent Similarity: 48.899 Percent Identity: 19.604

Alignment Block:

US-09-303-518D-125 x S39893 ..

Align seg 1/1 to: S39893 from: 1 to: 519

```

70 CCGGCGCATTCAGCGAGTCGGTTCGTCGCGAAGATATGCGCGATTCG 119
   |||||: : : : : ||| : : : : : ||| : :
42 ProAlaLeuIleArgLeuProLeu.....GlnGlnHisIleGlyAlaG1 56
   : : : : : |||||: : : : : ||| : : : : :
120 CCCCTCGATGAATCAAGAGAAAGCGGATCGCTCAAAAAGCCAGTGC 169
   : : : : : |||||: : : : : ||| : : : : :
56 uAlaGluProIleValIlyAspAspLeuValLeuIlyscIyGlnLeuI 73
170 TGTTCGAGACAAAGAAATCCGGCGTGTTCCTACTGCGCGCGCTCA 219
   : : : : : |||||: : : : : ||| : : : : :
73 LeAlaIlyAlaArgGlyProLeuSerAlaAsnIleHisAlaProThrSer 89
220 GCGAAATCGCGCGGAT.....CACGTCGCGAANA 251
   |||||: : : : : ||| : : : : :
90 GlyArgValIleAlaValGlyHisPheValAlaProHisAlaSerGlyLe 106
252 GCGCGTACTTCAGTCAGTCGTGATTCGCGTGAAGCAGACGANA.... 297
   ||| : : : : : ||| : : : : :
106 uProVal...ProThrIleThrIleArgProAspIlyGluAspIlyTrpG 122
298 ..ATGAGTTTCAGCCTACGACCTGAAGCGCTGCGAAGACTTAAAGCGC 345
   : : : : : ||| : : : : :
122 LyrProHisLeuProArgLeuArgProGlu.....AsnAlaAlaPro 135
346 GAGAAGTGGCGCGCGACCTGATCCATCCGATTTCG..... 384
   |||||: : : : : ||| : : : : :
136 GluGluIleAlaIleAlaGlnAlaAlaIleValIleValIleMetGlyG1 152
395 .....ACTGGCTGGCGACCGCTCCGCTTCAGCAAAATTCCTG 421
   : : : : : |||||: : : : : ||| : : : : :
152 yAlaThrPheProSerAlaValIlyLeuAsnLeuArgAlaIlyAspI 169
422 CCGTCGATGCGGAGCGCTTCGCAATCTTCGCAATGCGATGAGACACCAAT 471
   : : : : : |||||: : : : : ||| : : : : :
169 euThr.....ThrIleIleIleAsnGlyAlaGlyIleGlyGln 480
472 CCGCTGCTGCCGACCTAGCGTATTCATTAAGAAAGCCGCGAGATTT 521
   ||| : : : : : ||| : : : : :
181 ProTyrLeuThrCysAspAspArgLeuMetArgGluArgAlaGluGlu 197
522 CAAGCGCGCGCTGTGATTCGAGCCCTTTGACGAGCAAGCAAAATCCATG 571
   : : : : : |||||: : : : : ||| : : : : :
197 euAlaAspIlyIleGlyIleMetAlaAlaArgAlaLeu..GlyValIlyGlnVal 213
572 TTTGTA..... 577
214 PheValAlaIleGluSerAsnIlySerProGlnAlaIleGluAlaMetThr 230
578 .....AGCAGCTGGCGAGACGCTGCGCTGAAAAATGCTGCCAAT 620
230 gTyrAsnArgAlaLeuGlyTyrThrPheIlyHisValValProThrG 247
621 CGAAACACATGAATTCGCGCGCGCGCGATCCCTGCGTTTGATGGACACG 670
247 InTyrProMetGlySerGlyIlyHisLeuValIlyMetIleThrGlyGln 263
671 ACATTCATTCATGAGCGCGCGCGCGATTAACCGCTGCGACATC 720
264 GluThrProAlaArgAlaLeuThrAlaAspLeu..GlyValValHisA 280
721 AATATACAGATGTAATTCATTCATTCATTCATTCATTCATTCATTCAT 770
280 snIleAlaThrAlaHis..AlaValHisLeuAlaValArgTyrGlyGluPr 296
771 GAACACCGAGCGCGTATTCGCTAGGTTCATCAAGTCAACAAACCGC 820

```

```

296 OleuIleAlaArgThrValThrValSerGlyHisGlyIleArgArgPro 313
821 GCCCTTCGCGTACCGCTTTGGTCGCGAAGTATCGAATTAATTCGCG... 867
313 LaAsnLeuArgValIleuIleGlyThrProValSerIleuIleAlaHis 329
868 ...GCGCAATTCGTTGACACAGCAACCGCGTATTCGCTTCGCTAT 914
330 CysGlyIlyPheThrGluGluProAspArgLeuLeuGlyIlyProkme 346
915 GAACGCGCGGATTCACACAGCGCGCGACGATTTTGGACGCTACACA 964
346 tMetGlyMetProIleGln.....A 353
965 ATCAGATTCCTTCGTTATCGAAGAGCGCGACGCAAGAGCTGTTGCTGG 1014
353 snProArgValProValValIlyGlyThrAsnIlyIleuAlaLeuThr 369
1015 GTTCGCGCGCGAGCGGCAAAATTCATCCATCAGCGTACAAACCTCGGCA 1064
370 AlaAlaGluThrProGluAlaIlyThrMetProCysIleArgCysIlyAr 386
1065 TTTTCCTGAACAAACACTTCAGTTCAACACAGCGCTCAACGCGCGG 1114
386 g.....CysValGlnGly.... 390
1115 ACCGCGCATGTCGCGATTCGCTTACGAGCGCGTATGCGCTTGAT 1164
391 .....CysProValGly 394
1165 ATCTGCGCCACCTGCTTTGCGCGATTAATTCGCGGATACGACAG 1214
395 LeuThrProGluIleuAsnAlaArgIleHisAlaGlyAspLeuGlu 411
1215 CCGCGACGCGATTCGCTTCGTTGCAATTCGCGAAGAGAC 1255
411 yAlaAlaIlyValIlyLeuMetAspCysLeuAla.....C 423
1256 TCGCTTTGTCAGCTTCGTCGTCGCGGCAATTCAGATACGCGCGCTG 1305
423 yslIlyCysCysSerTyrAsnCysProAlaAsnLeu.....ProLeu 436
1306 TTGCGC 1311
437 ValGln 438
seq_name: p1r2:AE0692
seq_documentation_block:
probable NADH reducing dehydrogenase STY1665 [imported] - Salmonella enterica subsp.
C:Species: Salmonella enterica subsp. enterica serovar Typh
A:Note: this species has also been called salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
C:Accession: AE0692
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr
/ S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se
A:Reference number: AB0502; PMID:11677608
A:Accession: AE0692
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-673 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD01910.1; PID:q16502752; GSPDB:GN00176
A:Gene: STY1665
C:Superfamily: unassigned ferredoxin 2[4Fe-4S]-related proteins; ferredoxin 2[4Fe-4S]
alignment_scores:

```

Quality: 144.50 Length: 458
 Ratio: 0.698 Gaps: 18
 Percent Similarity: 45.197 Percent Identity: 21.616

Alignment_block:
 US-09-303-518D-125 x AE0692 ..

Align seg 1/1 to: AE0692 from: 1 to: 673

```

100 GAGAGATATGCGCGGTATGCGCCCTCGATGAAGTCAAGAGCGCATGC 149
    : : : : : : : : : : : : : : : : : : : : : : : :
48 LysGlnHisIleGlyAlaGluGlyIleuCyValSerValGlyAspAr 64
150 CGTCAAAAAGAGCGAGTGTGTTGAAGCAAAAGATCCGGCGGTGG 199
    ||| : : : : : : : : : : : : : : : : : : : : :
64 gvalIeuArGlyGlnAlaIeuThrArgGlyArGlyArMetLeuProV 81
200 TGTTTACTGCGCGCGCTTCAGGCAAAATCGCGCGATP..... 237
    ||| : : : : : : : : : : : : : : : : : : : :
81 aHis...AlaProThrSerGlyThrValIleAlaIleAlaProHisSer 96
238 .....CACCGTGGCGAAAGCGCGTACTTCAGTCAGTCGTGATTCGGCT 281
    ||| : : : : : : : : : : : : : : : : : : : :
97 ThrAlaHisProSerAlaIeuAlaGluLeu...SerValIleIleAspAl 112
282 TGAAGCGCAAGCAG..... 294
    : : : : : : : : : : : : : : : : : : : :
112 aspPolYeuIuAspArgTrpIleGluArGluGlyTrpSerAspYrArG 129
295 .....GAAATCGAGTTTGAACGCTAGCGACCTGAAAGCGGTGCA 333
    ||| : : : : : : : : : : : : : : : : : : : :
129 IahIserArgGluAlaIeuIleGluArGluIleHisGluTrGlyAla 145
334 AACTTAAAGCGGGAAGAAAGTGGCGCGCAACCTGATCCATCCGCTTGTG 383
    : : : : : : : : : : : : : : : : : : : :
146 gIyLeuGlyGly.....AlaGlyPheAr 153
384 GAGTGGCGTGGCGCACCGCTCGTCAGCAAAATTCCTGCGCGGTGGCG 433
    ||| : : : : : : : : : : : : : : : : : : : :
153 oThrGlyAla.....LysIleuGlnIleGlyAspArPL 164
434 AGCGGTTCGCCATCTCTGCAATGCGATGAGACCAATCCGTGGTGGC 483
    : : : : : : : : : : : : : : : : : : : :
164 yslIeThrThrIeuIleIleAsnAlaIaGluGlyProTyrlIleAr 180
484 GACCTCAGCGTATATCAAAAGACCGCGGAGATTCAACAGCGGGCT 533
    : : : : : : : : : : : : : : : : : : : :
181 AlaAspAspArgLeuMetGlnAspCysAlaAlaGlnIleValGluGly 197
534 GTTGGTATGAGCGGTTGACCGAA..... 558
    : : : : : : : : : : : : : : : : : : : :
197 eArgIleuAlaHisIleIleuGlnProArgGluValIleuIleGlyIleG 214
559 .....CGCAAAATCCATGTTTGAAGCGACGTCGCGCAGAC 594
    : : : : : : : : : : : : : : : : : : : :
214 IuAspAsnIysProGlnAlaIleSerMetLeuArgAlaValIleuAlaAsp 230
595 GTGCCGTGCAAAATGCTGCCAATCGCAACATGAATTCGGCGGGCC 644
    ||| : : : : : : : : : : : : : : : : : : : :
231 .....AlaHisAspIleSerLeuArGValIleProThrly 242
645 GCATCTGCGCGGT.....TTGAGTGGCAGCG 670
    : : : : : : : : : : : : : : : : : : : :
242 sTyfProSerGlyGlyAlaLysGlnIleuThrGlnIleuThrGlyLys 259
671 ACATTCATTTTCATCGAGCGGTGCGCGCAATAAACCGTGGGACCATC 720
    : : : : : : : : : : : : : : : : : : : :
259 IuVal.....ProHisGlyGlyArgSerSerAspIleGlyVal 271
721 AATTATCAAGATGTAATTAC.....ATTGCGCGTTGTTCGCAAC 761
    ||| : : : : : : : : : : : : : : : : : : : :
272 LeuMetGlnAsnValGlyThrAlaTyAlaValIysArgAlaValAlas 288
  
```

```

762 AGGCGGTTCGAACACCGAGCGGTGATTCGCTAGGTGTTCTCAATCA 811
    ||| : : : : : : : : : : : : : : : : : : : :
288 pGlyGluProIleThrGluArgValAlaThrIleuThrGlyAlaIle 305
812 ACNAACCGCGCGCTTTCGCTACCGTTTGGGTGCGAAAGTATCGCAAT 861
    : : : : : : : : : : : : : : : : : : : :
305 eArArgProGlyAsnValItrpAlaArgIleuGlyThrProValArgHis 321
862 .....ACTGCGGGCGCAATGTGTGACACAGACAAACCGGTGATTTCG 905
    ||| : : : : : : : : : : : : : : : : : : : :
322 LeuAsnAspAlaGlyPheCysProSerAlaIaAspGlnMetValIle 338
906 TTCGTGATTTCGAAGCGCGCATTCACAGAGCGCGCAGATATTTCGGAC 955
    ||| : : : : : : : : : : : : : : : : : : : :
338 yGlyProIleuMetGlyPheThrIleuProItrpIleAspValProVal 355
956 GCTACCAACATCAGATT.....TTCGTTATCGAAGAGCGCGCGCAC 996
    : : : : : : : : : : : : : : : : : : : :
355 yslIeThrAsnCysIleuIleAlaProSerValThrGluMetGly..... 369
997 AAAGAGCTGTTCGGGTGGTGGCGCGCAGCGGCAAAATCTCATCTCAC 1046
    ||| : : : : : : : : : : : : : : : : : : : :
370 .....AlaProGlnIleGluIlySerCysIleAr 379
1047 GCGTACAAACCTCGCCCATTTCTGAAAAACAAACTTTCATGTTACAA 1096
    : : : : : : : : : : : : : : : : : : : :
379 gCysSerAlaCysAlaAsp..... 385
1097 CAGCGGTCAACGCGCGCGAGCGGCGCATGTGCGCATTTGTTACTTAC 1146
385 ..... 385
1147 CGCGTATGCGCTTGATATCTGCGCACCGCTCTTTGGCGGATTTAAT 1196
    : : : : : : : : : : : : : : : : : : : :
386 ...AlaCysProAlaAspIleuProGlnIleuThrIleuThrPheSer 401
1197 CGTGGCGATTCGACAGCGCGCAGGCA.....TTCGGTTCCTTGG 1237
    : : : : : : : : : : : : : : : : : : : :
401 sGlyGlnGlnHisAspYsaIatThrAlaHisIleAlaAspCysIleG 418
1238 AATTGAGCAAGAAAGACCTGCTTGGCAGCTTCGTCGCGCGGCGAA 1287
    ||| : : : : : : : : : : : : : : : : : : : :
418 Iu.....CysGlyAlaCysAlaItrpValCysProSerAsn 429
1288 TACGATACGCGCGCTGTTCGCC 1311
    ||| : : : : : : : : : : : : : : : : : : : :
430 Ile.....ProIleuValGln 434
  
```

seq_name: pIrf2:E72398

seq_documentation_block:

hypoethetical protein TM0244 - Thermotoga maritima (strain MSB8)
 C:Species: Thermotoga maritima
 C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
 C:Accession: E72398
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, J.C.; Garret, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, C.M.
 Nature 399, 323-329, 1999
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome
 A:Reference number: A72200; M01D:99287316
 A:Accession: E72398
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1451 <ARN>
 A:Cross-references: GB:AE001708; GB:AE000512; NID:94980740; PIDN:AAD5335.1; PTD:9498
 A:Experimental source: strain MSB8
 C:Genetics:
 A:Gene: TM0244

alignment_scores: Quality: 140.00 Length: 454

Ratio: 0.648 Gaps: 17
Percent Similarity: 47.577 Percent Identity: 20.264

alignment_block:

US-09-303-518D-125 x E72398

Align seg 1/1 to: E72398 from: 1 to: 451

```

49 GAGCAAGCCGTTTACGACGCGCCGATTCAGCAAGTCGCGTTCGTTGG 98
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
35 AspLysProIleGluValArgLalProLeuProGlnLysValPheValPheLe 51
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
99 GCAAAATATGCGCGGTATGCGCCCTCGATGAAATCAAGAGGCGCATG 148
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
51 uSerSnhSalaGlaYasnProAlaLysProValSerProGlyAspG 68
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
149 CGGTCAAAAAGGCCAGTGTGTGTAAGAAGAAAGATCCGGCGCTG 198
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
68 LuValLysThrGlyGlnValIleGlyGluProGluGlyPheIleSerAla 84
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
199 GTGTTTACTGCGCGCGCTTCAGGCAAAATCGCCGATTCACCGTGCGCA 248
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
85 TyrLeuHISerProValThrGlyArgValLeuIle.....GuileI 97
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
249 AAAGCGCGTACT.....CAGTCAGTCGCA 274
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
98 .LysLysIleLeuHISProIleLeuGlyLysProIleGluAlaIleVal 114
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
275 TTGCGGTGAAGGCAAGCAGCAAAATC.....GAGTTT 306
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
114 LeGluArgThrSerAspSgluTrpValHISIleGluThrGlyAspPhe 130
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
307 GAGCGCTACGACCTGGAAGCGCTGGCAACTTAAGCGGCAAGAGTGGC 356
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
131 GluArgMetSerLysGluGluIleLeu.....GuileI 142
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
357 CCGCAACCTGATCCAAATCGGTTGTGAGTACGCTGGCAACCGCTCCGT 406
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
142 eLysLysAlaGlyIleValGlyLeuGlyGlyAlaMetPheProThrHisV 159
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
407 TCAGCAAAATTCCTGCGCGTGCATGCGAGCGCTCGCCATCTTCGTCAAT 456
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
159 aLysLeuSerProProGluLysLysValAspThrLeuIleValAsn 175
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
457 GCGATGACACCAATTCGCTGGCGTCCGACCCCTACGCTCATTCAAAGA 506
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
176 GlyAlaGluCysGluProValLeuThrIleAspHisArgLeuMetLeuG 192
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
507 AGCGCGCAGAGATTTCAAAGCGCGCTGTGTGATTGAGCGCTTGACCG 556
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
192 uArgLalGluAspIleLeuGlnGlyIleLeuIleMetLeuVal..... 207
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
557 AACCCAAATTCATGTTTGAAGCA...CGTGGCGCAGACGTCGCTCT 603
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
208 .....LeuGlyValGlnLysAlaValGlyValGlnSerAsnLys 221
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
604 GAAATCTCTGCAACATGAAACACATGATGCGGGCGCCGATCTCTG 653
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
222 MetAspLalArgHisAsnLeuLysValPheLysGly...TyrProVa 237
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
654 CGGTTGAGTGGCAGCAGCATTCATTTCATCGAGCGGTCGCGCGAANA 703
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
237 lAspValAlaLeuLeuArgThrLysTyr.....ProGlnGlyAlaGlu 252
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
704 AA...ACCGTGGGACCATC..... 720
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
252 ySgluLeuIleTyrAlaIleThrGlyArgMetValProArgLysGlyLeu 268
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
721 .....AATTCAGAGATGATTAACAT 743
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
269 PrometAspValGlyValValValGlnAsnValGlyThrCysValAlaVa 285
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
744 TGCGCGTTGTTTGCACAGCGCGTCTGAAACACCGAGCGCGTGAATGGCC 793

```

```

285 lLysGluAlaValValAspIleLysProLeuValGluArgLysMetThrV 302
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
794 TAGGTGTTCTCAAGTCACAAACCGCGCTTCGTGCTACCGTTTGGT 843
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
302 aISerGlyAspAlaValLysAsnGlnLysAsnLeuIleValArgIleGly 318
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
844 GCGAAATATCGCAATTACT.....GCGGCGAAATGGTTGACACAGA 887
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
319 ThrProValLysAspAlaIleAspTyrCysGlyGlyIleAspLysSnh 335
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
888 CAACCGCGATTCGCGTTCGCTGATTCGAACGC...GGATTACGCAAG 934
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
335 rGluArgValIleLeuGlnGlyProMetMetGlyIleSerIleThr... 350
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
935 GCGCGCAGCATTAATTGGGACGCTACCAATCAATGATTCGTTATCGAA 984
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
351 .....AsnLeuAspIleProValMet 357
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
985 GAAGCGCGCAGCAAGAGCTGTTGCGCTGGCTTGGCGCCGACCGGACAA 1034
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
358 LysGlyThrSer..... 361
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
1035 ATACTCCATCACGCGTACACCCCTCGCCATTCCTGAAAACAAACTCT 1084
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
362 ....GlyIleThrAla.....PheLeuProLys... 370
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
1085 TCAGTTCAACACAGCGCTCAACGGCGGCGCGCCGCTGTCGCATTT 1134
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
371 .....SerArgProGlnLysProCysIleArgCys 380
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
1135 GGTACTTACGAGCGCTGATGCCCTTGATATCCGATCCGCCACCTGCTTT 1184
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
381 SerGluCysValGlnValCysProMetAsnLeuGlnProTyrLeuLeuT 397
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
1185 GCGCATTTAATCGTGGGATACGACAGCGCGCAGCATGCTGCTTCT 1234
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
397 rLeuLeuSerThrLysArgLysTyrAspGluLalValGluAsnGlyLeuM 414
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
1235 TGGAAATTGAGCAGAGAAGACCTCGCTTGTGACGTTGCTGCGCCGCGC 1284
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
414 eAspCysIleGlu.....CysGlySerCysThrTyrThrCysProSer 428
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
1285 AATATCAGATAC 1296
   |||  |||||
429 lLysIleGluHIS 432

```

seq_name: p1r2:E64136

seq_documentation_block:

rnfc protein homolog - Haemophilus influenzae (strain Rd KW20)

C/Species: Haemophilus influenzae

C/Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #extl_change 17-Mar-2000

C/Accession: E64136

R:Flieschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, J.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M. Science 269, 496-512, 1995

A:Authors: Guelm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.

A:Reference number: A64000; MUID:95350650

A:Accession: E64136

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-819 <TIGR>

A:Cross-references: GB:U32841; GB:I42023; MID:q1574529; PIDN:AAC2331.1; PID:q1574537

C:Superfamily: unassigned ferredoxin 2[4Fe-4S]-related proteins; ferredoxin 2[4Fe-4S] F;371-435/Domain: ferredoxin 2[4Fe-4S] homology <FER>

alignment_scores:

Quality: 138.50 Length: 468
Ratio: 0.635 Gaps: 17

Percent Similarity: 46.581 Percent Identity: 19.658

Alignment block:

US-09-303-518d-125 x E64136 ..

Align seg 1/1 to: E64136 from: 1 to: 819

```

70 CCGGCATTACCGAAGTCGGCTGCTGGCCAGAAATATGCCGGATGCG 119
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
38 ProLeuGlyThrAspPheTyrIleProLeuLysGlnHisLeuGlyThrTh 54
120 CCCCTCGATGAAATGCAAGAGAGCGGATGCCGTCAAAAAGGCGCAATG 169
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
54 rGlyAsnLeuLeuIleLysGlnGlyAspTyrValLeuLysGlnIleAla 71
170 TGTTCGAGACAAAAAGATCCGGCGGTGTGTACTTACCGCGCGCTTCA 219
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
71 euThrLysGlyAspLysLeuArgMetLeuProValHisAlaProThrSer 87
220 GCGAAATGCGCGCGATTCACCGTGGCAAAAGCGGTAATCTTCAATCAGT 269
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
88 GlyThrIle.....LysSerIleLysProLysTyrVal 97
270 CGTGATGCGCGTGAAGCAAGCAAGAAATCGAG..... 303
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
97 lAlaThrHisProSerGlyLeuAspLysProThrIleHisLeuGlnAla 114
304 .....TTGAACGCTACGACACTGAAGCGCTGGCAAC 336
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
114 spGlyLeuAspLysProIleGlnArgAsnProIleAspSerPheSerThr 130
337 TTAAGCGGCAAGAGATGCGCGCAACATGATCAATCCGTTGTGAC 386
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
131 LeuSerSerGlnGlnLeuIleHisLysIleTyrGlnAlaGlyLeu...Al 146
387 TGGCGTGGCGACCCGTCGCTC.....AGCAAAATTCCTGGCGTGC 427
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
146 agLysLeuGlyGlyAlaValAlaPheProThrAlaAlaLysIleGlnSerAla 163
428 ATGCGAGCGCGTTCGCCATCTTCGTCATGATGAGACACCAATCCGCTG 477
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
163 LysGlnLysValLysLeuLeuIleIleAsnGlyAlaGlnLysGlnProTyr 179
478 GCTGCGGACCTTACGCTCATTTATCAAGAAGCCCGGAGGATTTCAACG 527
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
180 lIleThrCysAspAspArgLeuMetArgGlnArgAlaAspLysIleLeu 196
528 CGGCGTGTGGTATTTGAGCGGTTTGACGACGCAAAATCCATGTTGTA 577
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
196 sGlyIleArgIleLeuArgTyrIleLeuHisProGlu...LysValValI 212
578 AGGCACTGGCGCAGACGTCCGCT..... 603
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
212 lAlaIleGlnLysAspAsnLysProGlnAlaIleSerAlaIleArgAsnAla 228
604 ...GAAATGCTCCACATCAACATCAATGATTTGGCGGCGCGCATGC 650
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
229 LeuGlnGlyAlaAsnAspIleSerIleArgValIleProThrLysTyrPr 245
651 TGGCGGTTGAGTGGCAGCAGCATTCATTTCAATCAGAG.....C 688
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
245 oSerGlyAlaThrLysGlnLeuIleTyrIleLeuThrArgLysIleGlnValP 262
689 CGGTCGCGCGGAATAAACCGTGTGACCATCAATTAATCAAGATTAAT 738
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
262 roSerGlyLysArgSerSerIleGlyValLeuMetGlnAsnVal... 277
739 ACATTTGGCGGTTGTGTTGCAAGGCGCT.....CTGAC..... 774
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
278 .....GlyThrMetPheAlaIleLysArgAlaIleLysAspLysP 292
775 .....ACGACGCGGCTATTCAGTGGTGTCTCAAGTCAACAAACCCG 820
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

292 oLeuIleGlnArgValValThrLeuThrGlyAsnLysIleAlaGlnLysG 309
821 GCCCTTGCGTACCGTTTGGTGGCGAAAGTATCGCAAT.....ACT 864
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
309 LysAsnTyrTrpValArgLeuGlyThrProIleSerGlnIleLeuSerAsp 325
865 GCGGCGCAATTTGGTGTGACACAGAACACCGCGGATTTCCGCTGAT 914
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
326 AlaGlyTyrGlnPheAspLysHisPheProIlePheAlaGlyLysPro 342
915 GAAAGCGCGCATTAACACAGCGCGCAGCATTAATTTGGACGTACCA 964
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
342 tMetGlyLeu..... 345
965 ATCAGATTTCCGTTATCGAAGAGCGCGCAAGAGCTGTGCTGCTG 1014
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
346 ..GluLeuProAsnLeuAsnAlaProValThrLysLeuValAsnCysLeu 361
1015 GTTGGCGCGGACCGGCAAAATCTCCATCAGCGGTACAAACCTCGGCA 1064
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
362 LeuAlaProAspTyrLeuGlnTyr..... 369
1065 TTTCCTGAAAAACAACCTTCAAGTTCAACACAGCGGTCAAGCGCGG 1114
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
370 .....AlaGlnProGlnAlaG 375
1115 ACCGCGCCAGTGGCGCATTTGACTTACAGAGCGCGGTGATCCCTTGAT 1164
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
375 LysGlnAlaCysIleArgCysSerSerCysSerAspAlaCysProValAsn 391
1165 ATCCTGCGCCACCCGCTTGGCGGATTTAATGCTGCGCATACCGACAG 1214
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
392 LeuMetProGlnGlnLeuTyrTrpPheAlaArgSerLysAspHisLysI 408
1215 CGCGCAGCGCATTTGGT.....TCCTTGGAATTTGACGAAAGAACACC 1255
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
408 sSerGlnGlnTyrAlaLeuLysAspCysIleGln.....C 420
1256 TCGCTTTGTCAGCTTGCTGCTGCGCGGCAATTCGATACGCGCGCTG 1305
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
420 ySgIlyIleCysAlaTyrValCysPro.....SerHisIleProLeu 433
1306 TTGCGC.....AAAGTCTGGAACCATTTGAA 1334
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
434 lIleGlnTyrPheArgGlnGlnLysAlaLysIleTrpGlnIleLysGln 450
1335 GGA 1338
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
450 sGln 451
seq_name: p12:569703
seq_documentation_block:
HKRI protein precursor - yeast (Saccharomyces cerevisiae)
N:Alternate names: Protein YDR420w
C:Species: Saccharomyces cerevisiae
C>Date: 22-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 23-Mar-2001
C:Accession: S69703; A53382
R:Dietrich, F.S.
submitted to the EMBL Data Library, August 1995
A:Description: The sequence of S. cerevisiae lambda 3641 and cosmids 9461, 9831, and
A:Reference number: S69555
A:Accession: S69703
A:Molecule type: DNA
A:Residues: 1-1802 <DIE>
A:Cross-references: EMBL:033007; NID:9927685; PIDN:AAB64857.1; PID:9927691; MIPS:YDR4
R:Kasahara, S.; Yamada, H.; Mio, T.; Shitatori, Y.; Miyamoto, C.; Yabe, T.; Nakajima,
J. Bacteriol. 176, 1488-1499, 1994
A:Title: Cloning of the Saccharomyces cerevisiae gene whose overexpression overcomes
A:Reference number: A53382; MIMD:94156857
A:Accession: A53382
A:Molecule type: DNA
A:Residues: 1-581, 'A', 583-593, 'A', 595-1802 <KAS>

```

A:Cross-references: EMBL:S69101; NID:9545659; PIDN:AB30051.1; PID:9545660
 A:Experimental source: YNN25
 A:Note: sequence extracted from NCBI backbone (NCBIN:144410, NCBI:P:144411)
 C:Genetics:
 A:Gene: SGD:HKR1
 A:Cross-references: SGD:S0002828; MIPS:YDR420W
 A:Map position: 4R
 C:Keywords: calcium binding; glycoprotein; transmembrane protein
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-1802/Product: HKR1 protein #status predicted <MAT>
 F:1483-1508/Domain: transmembrane #status predicted <TM>
 F:1645-1656/Domain: calcium binding #status predicted <CAL>

alignment_scores:
 Quality: 135.00 Length: 497
 Ratio: 0.560 Gaps: 22
 Percent Similarity: 48.491 Percent Identity: 22.736

alignment_block:
 US-09-303-518d-125/rev x 569703 ..

Align seg 1/1 to: 569703 from: 1 to: 1802

```

1337 TCCCTCTCAATGTTTCCAGCAGCTTGGCAACAGCGCGGATTCGTA 1288
|||||.....|.....|.....|.....|.....|.....|.....|
390 SerAspThrValSerThrSerThrAsnThrValProTyrSer 406
1287 TTGGCCGGGCGACGACGAGCTGCACAAAGCGAGCTCTTCTGTCGCAAT 1238
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
406 rVal...HisSerThrPheValHis...AlaThrSerSerThrTyrI 421
1237 CCAAGCAACCCAAATGCTGCGCGCTGCGATCGCCGACGATTAATG 1188
.....|.....|.....|.....|.....|.....|.....|
421 IeSerSerSerLeuTyrSerSerProSerLeuSerAlaSerValSer 437
1187 GCGCAAAAGCGGTGGCGACGATATCCAAAGCATCAGCGCTGTAACT 1138
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
438 HisPheGlyValAlaProPheProSerAlaTyrIleSerPheSerVal 454
1137 ACCAATC.....GACCATGCGCGCGCGCGCTGTCGAGCGGTG 1097
|||||::|::|::|::|::|::|::|::|::|::|::|::|::|
454 IProValAlaValSerSerThrTyrThrSerSerProSerAlaSerVal 471
1096 TGTGACACTGAAGATTGTTTTCAGGAATGCGGAGGTGTACGC 1047
||.....|.....|.....|.....|.....|.....|.....|
471 al.....|.....|.....|.....|.....|.....|.....|
1046 GTGATGAGATTTTTCGCGCTGCGCGGCAACCGCGAAGCTCTT 997
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
472 .....ValProSerAlaTyrAlaSerSerProSerValProva 484
996 GCTGCGGCTCTTCGATACGGAATCTGATGTGAGGTCGCAAT 947
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
484 lAlaValSerSerThrTyrThrSer.....|.....|.....| 493
946 AATGTCGCGCGCTGTGTAATCGCGCGTCAATACGAACCGGAATC 897
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
494 ..ProSerAlaProAlaIleSerSerThrTyrThrSerProSer 509
896 ACGGCGTTGTGTGTCAC.....AATTCGCGCGC 865
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
510 AlaProValAlaValSerSerThrTyrThrSerProSerAlaPro 526
864 AGTAATTTGGCATCTTC.....GCACCAAAAACGGTAC 830
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
526 aAlaIleSerSerThrTyrThrSerProSerAlaProValAlaVal 542
829 GCAAGAGCGCGGTTTGTGACTTGAGACCA.....CCTAGGCAATC 786
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
543 .....SerSerThrTyrThrSerSerProSerAlaProAlaIle 556
785 ACGGCTCGGTGTTCAGACGCGCTGTTCAAACAAACGGCAATGTAAT 736

```

```

557 SerSerThrTyrThrSerSerProSerAla.....ProValAlaVal 570
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
735 TACATCTTGATTAATGATGTCACACAGCGTTTATTCGCGCGACGCGC 687
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
570 IeSerSer.....ThrTyrThrSerSerProSerAla 581
687 .....|.....|.....|.....|.....|.....|.....| 687
581 roValAlaIleSerSerThrTyrThrSerSerProSerValProValAla 597
686 ...TCGATGAAGATGTCGTCACCAATCAACCGCAGAGA..... 648
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
598 ValSerSerThrTyrThrSerSerProSerAlaProValAlaIleSer 614
647 .....TCGCGCGCGCGCAATTCATGATGTTTCAGATTCGACGATTT 605
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
614 rThrTyrThrSerSerProSerAlaProValAlaValSerSerThrTyr 631
604 CAGACGCGACGTCGCGCGAGCTGCTTACCAACATGATTTGCGTTCG 555
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
631 hrSerSerProSerAlaProAlaIleSerSerThrTyrThrSer 647
554 .....GTCAACGCGCTCAATTCACCAACAGCGCGCTTT 523
648 ProSerValProValAlaValSerSerThrTyrThrSerPro..... 662
522 GAATCTCTCGCGGCTCTTGTATATGACCCGAGGCTGGCAGCCAGCG 473
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
663 ....SerAlaProAlaIleSerSerThrTyrThrSerSerProSer 678
472 GATTGTGTCATGCGATGACGAAGATGCGAAGCGCGTCGACATGACG 423
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
678 alProValAlaValSerSerThrTyrThrSerSerProSerAlaPro 694
422 GCAAGATTTTCTGAACGAGCGGTGCGCAGCGCATCCACAAACCGCA 373
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
695 Ala.....IleSerSerThrTyrThrSerSerProSerAlaPro 708
372 TTGATCAGAGTTGCGCGCAGCTCTTCGCGGCTTAAGTTGCCAGCGCT 323
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
708 lAlaValSerSerThrTyrThrSerSerPro.....SerAla 721
322 CAGGTGCGTAGCGCTTCACATCGATTTGCTGTCGTCCTCAACGCGCAT 273
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
721 roAlaIle...IleSerSerThrTyrThrSerSerProSerAlaProVal 736
272 ACCACTGACTGAAGTACGCGCTTTTCGCCACGATGATTCGCGATTTT 223
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
737 AlaValSerSerThrTyrThrSerSerProSerAlaProAlaIle... 752
222 GCCTGAAGCGCGCGCAGTAAACACCAACGCGCGGATTTTGTCTTCA 173
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
753 .....SerSerThrTyrThrSerSerProSerAlaProValAlaVal 767
172 ACAGCACTTGCGCTTTTTCAGCGCATGCGCTTCCTGACTTCATTCAG 123
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
767 erSerThrTyr.....ThrSerSerProSerAlaLeuValValLeu 780
122 GAGCGCATACGCGCATATCT.....TCGCCAAGCA 91
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
781 SerSerThrSerThrSerSerProTyrAspIleValTyrSerProSer 797
90 CCGCATTCGATATGCGCGG.....CCGTCGTAACCGCTTCCTCG 47
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
797 rPheAlaIleIleSerSerGlyTyrThrProSerProSerAlaSerVal 814
46 GTCTGCGCGCATGGGAGGTTTAGACCTTTTATTATTA 6
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
814 lMetSerSerThrSerSerSerProTyrAspIleVal 827

```

seq_name: pIir2:TO2345

seq_documentation_block:
 hypothetical protein KIA0324 - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 05-Nov-1999
 C:Accession: T02345
 R:Ricke, D.O.; Bruce, D.; Mundt, M.; Doggett, N.; Munk, C.; Saunders, E.; Robinson, D.;
 re, J.; White, S.; Ueng, S.; Tatum, O.; Campbell, C.; Fawcett, J.; Deaven, L.
 submitted to the EMBL Data Library, March 1998
 A:Description: Sequencing of human chromosome 16p13.3.
 A:Reference number: Z14664
 A:Accession: T02345
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1791 <RIC>
 A:Cross-references: EMBL:AC004493; NID:g2996648; PIDN:AAC08453.1; PID:g2996650
 C:Genetics:
 A:Map position: 16
 A:Introns: 1610/2; 1706/2
 A:Note: KIA0324

alignment_scores:
 Quality: 131.00 Length: 411
 Ratio: 0.697 Gaps: 14
 Percent Similarity: 45.742 Percent Identity: 25.061

alignment_block:
 US-09-303-518D-125 x T02345

Align seg 1/1 to: T02345 from: 1 to: 1791

```

12 CAAAAGAGTCTAAACCTGCCATCGCGGAGACCGGACCAAGCCGTTT 61
   ::::|||||::: |||::: |||::: |||
776 ArgArgArgSerAlaSerSerProArgThrLysThrThrSerArgGly 792
   :::: |||::: |||::: |||::: |||
62 ACGACGGCC.....GCCATTCACGAAGTCGCGTT.... 92
   ||| ||| ||| ||| ||| ||| ||| ||| |||
792 YArgSerProSerProLysProArgGlyLeuGlnArgSerArgSerArg 809
   :::: |||::: |||::: |||::: |||
93 .....GCTGGGGAAGAAATATGCCGTATGCGCCC 122
   :::: |||::: |||::: |||::: |||
809 eArgArgGlyLysThrArgThrArgArgArgAspArgSerGlySer 825
   :::: |||::: |||::: |||::: |||
123 CTCGATGAAGTCAAGAGCGATGCCCTCAAAAAGGCCAAGTCGT 172
   :::: |||::: |||::: |||::: |||
826 SerGlnSerThrSerArgArgGlnArgSerArgSerArgSerArg 842
   :::: |||::: |||::: |||::: |||
173 TTGAAGCAAAAAGATCCGGCGCTGTTTACTCGCGCGCTTCAGGC 222
   ||| ||| ||| ||| ||| ||| ||| ||| |||
842 LThrArgArgArgArgGlyGlySerGlyTyrHis..... 853
   :::: |||::: |||::: |||::: |||
223 AAAATCGCGCGGATTCACCGCGGAAAGGCGTACTTCAGTCAGTCGT 272
   :::: |||::: |||::: |||::: |||
854 ..SerArgSerProAlaArgGlnLysSerArgThrSer..... 867
   :::: |||::: |||::: |||::: |||
273 GATTGCGGTTGAAGCAAGCAAGAAATCGATTGAAGCTACGACCTG 322
   ||| ||| ||| ||| ||| ||| ||| ||| |||
868 .....ArgArgArgArgGlyArgSerArgThrProProth..... 879
   :::: |||::: |||::: |||::: |||
323 AAGCGCTGGCAACTTAAGCGGCGAAGATGCCCGCAACTGATCAA 372
   :::: |||::: |||::: |||::: |||
880 .....SerArgLysArgSerArgSerArgThrSerProAlaPro 893
   :::: |||::: |||::: |||::: |||
373 TCCGTTTGTGAGCTGCGCTGCGCACCGCGCTTCAGCAAAATCCCTGC 422
   :::: |||::: |||::: |||::: |||
893 rLysArgSerArgSerArgAlaSerProAlaThrHisArgArgSerArg 909
   :::: |||::: |||::: |||::: |||
423 CGTGCATGCCGAGCGGTCCGACATCTTGCAATGCGATGACCAATC 472
   ||| ||| ||| ||| ||| ||| ||| ||| |||
910 SerArgThrProLeuLysSerArgArg..... 919
   :::: |||::: |||::: |||::: |||
473 CGTGGCTGCCGACCTACGCTATTCAAAGACCGCGCGAGATTC 522
   :::: |||::: |||::: |||::: |||

```

```

920 .....SerArgSerArg..... 923
523 AAACGGCGCTGTGTGATTGAGCCGTTGACCGAAGCAAAATCCAGT 572
   ||| ||| ||| ||| ||| ||| ||| ||| |||
924 ..ThrSerProAlaSer..... 928
573 TTGTAGGACAGTGGCGGACGAGCTGCGCTGTGAATGCTGCCACATCG 622
   ||| ||| ||| ||| ||| ||| ||| ||| |||
929 .....ArgArgArgSer..... ArgSerArg 935
623 AAACACATGAATTCGGCGCGCCGATCTCGCGTTTGAGTGCAACGAC 672
   :::: |||::: |||::: |||::: |||
935 gThrSerValThrArgArgSerArgSerArgAlaSerProAlaSerA 952
673 ATTCATTTCATCGAGCC.....GTCGGCGGCAATRA 704
   :::: |||::: |||::: |||::: |||
952 rArgArgSerArgSerArgThrProProValThrArgArgSerArg 968
705 AACCGTGTGACCATCATTAATTAAGATGAATTAATTAATTAATTAAT 754
   :::: |||::: |||::: |||::: |||
969 SerArgThrProThrThrArgArgArgSerArgSerArgThrProPro 985
755 TTGCAACAGCGCGTCTGAACACCGGCGGATTCGCTAGTGATGCT 804
   :::: |||::: |||::: |||::: |||
985 LThrArgArgSerArgSerArgSerArgThrProProValThrArgArg 1002
805 CAAGTCAACAAACCGCGCTCTTGCGTACCGCTTTGG..... 841
   ||| ||| ||| ||| ||| ||| ||| ||| |||
1002 eArgSerArgThrSerProLysThrArgArgArgSerArgSerArgThr 1018
842 .....GTGCAAGATGATCAAAATTAATTAATTAATTAATTAATTA 885
   ||| ||| ||| ||| ||| ||| ||| ||| |||
1019 SerProValThrArgArgArgSerArgSerArgThrSerProValThr 1035
886 GACACACCGCGTATTCGCGTTCGATTAAGACGCGCGATTAACAAG 935
   :::: |||::: |||::: |||::: |||
1035 gArgArgSerArgSerArgThrSerProValThrArgArgSerArg 1052
936 CGGCGACG.....ATTATTGGAGCGTACCAATACAGATTCGCTTA 979
   :::: |||::: |||::: |||::: |||
1052 eArgThrProProAlaLysArgArgSerArgSerArgThrProLeu 1068
980 .....TCGAGAGAGCGCGCAGCAAGAGCTGTGCGTGGTTCGCGCG 1023
   ||| ||| ||| ||| ||| ||| ||| ||| |||
1069 LeuProArgLysArgSerArgSerArgSerProLeuAlaLysArgArg 1085
1024 CAGCCGCAAAATCTCCATCAACGCGTACA..... 1054
   ||| ||| ||| ||| ||| ||| ||| ||| |||
1085 gSerArgSerArgThrProArgThrAlaArgGlyLysArgSerLeuTh 1102
1055 .....CCCTCGGCGAATTCCTGAAAACAACTTCGAAGTTCACACAG 1099
   ||| ||| ||| ||| ||| ||| ||| ||| |||
1102 rSerProProAlaLysArgArgSerAlaSerGlySerSerSerasp 1118
1100 CCGTCAACGCGCGGACCGCGCCATGTCGCGCA 1132
   ||| ||| ||| ||| ||| ||| ||| ||| |||
1119 ArgSerArgSerAlaThrProProAlaThrArg 1129

```

seq_name: p1r2:A40670

seq_documentation_block:
 nuclear envelope protein POM 121 - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 05-Nov-1999
 C:Accession: A40670
 R:Hallberg, E.; Wozniak, R.W.; Blobel, G.
 J. Cell Biol. 122, 513-521, 1993
 A:Title: An integral membrane protein of the pore membrane domain of the nuclear envelope
 A:Reference number: A40670; MUID:93328754
 A:Accession: A40670
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1199 <HAL>

A:Cross-References: GB:221513; MUD:9396746; PIDN:CA979725.1; PID:9396747
 F:803-807,845-849,956-960,1010-1014,1047-1051,1076-1080/Region: pentapeptide motif (X-F-

alignment_scores:

Quality: 129.50 Length: 477
 Ratio: 0.568 Gaps: 19
 Percent Similarity: 47.799 Percent Identity: 21.384

alignment_block:

US-09-303-518D-125/rev x A40670 ..

Align seg 1/1 to: A40670 from: 1 to: 1199

```

1340 CCTCTCTCTCAATGTTTCCAGACACTTGGCGAACAGCGGGCGTATTC 1291
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
546 PropriophetThreuthrLeupProthrValGlyProAlaAlaSerProAl 562
    : : : : : : : : : : : : : : : : : : : : : : : : : :
1290 GTATTGGCCGGGCGAGCG.....AAGCTGCACAAAG 1259
    : : : : : : : : : : : : : : : : : : : : : : : : : :
562 aserLeupProAlaProSerSerAsnProLeuLeuGluSerLeuLysM 579
    : : : : : : : : : : : : : : : : : : : : : : : : : :
1258 CGAGGCTCTCT...TCGTCCAAATCCAGAACCCCAATGCTGGCGCTG 1212
    : : : : : : : : : : : : : : : : : : : : : : : : : :
579 etGlnGluSerProAlaProSerSerSerGluProGluAlaAlaThr 595
    : : : : : : : : : : : : : : : : : : : : : : : : : :
1211 TCGGTATGCGCGAGATTAAATCGCGCAAAAGCAGGCGGCGAGATATC 1162
    : : : : : : : : : : : : : : : : : : : : : : : : : :
596 VALAlaAlaProSerProProLysThrProSerLeuAlaAlaProLeuVa 612
    : : : : : : : : : : : : : : : : : : : : : : : : : :
1161 CAAGGGCATCAGCGCGCTGTAAGTACCAATC...GGCACCATGGCGCGGT 1115
    : : : : : : : : : : : : : : : : : : : : : : : : : :
612 IserProLeuThrGly.....ProLeuAlaSerThrSerSerAsps 626
    : : : : : : : : : : : : : : : : : : : : : : : : : :
1114 CGCCGCCCTGAGCGGCTGTGAACTGGAAGAT..... 1080
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
636 eLysProThrThrThrPheLeuGlyLeuAlaSerAlaSerSerAlaThr 642
    : : : : : : : : : : : : : : : : : : : : : : : : : :
1079 ...TTGTTTTCAGGAATGCGCGAGGTTGTACGCGGTGATGATATT 1033
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
643 ProLeuThrAspThrLysAlaAlaProGlyValSerGlnAlaGlnLeuCySva 659
    : : : : : : : : : : : : : : : : : : : : : : : : : :
1032 GTCCGGCTGGCGCGGCAACCCAGACGACGCTCT..... 999
    : : : : : : : : : : : : : : : : : : : : : : : : : :
659 IserThrProAlaAlaThrAlaProSerProThrProAlaSerThrLeup 676
    : : : : : : : : : : : : : : : : : : : : : : : : : :
998 .....TTGCTGGGCGCTTCTTCGATAACGGAATCTGATTGGTAG 957
    : : : : : : : : : : : : : : : : : : : : : : : : : :
676 heGlyMetLeuSerProProAlaSerSerSerSerLeu.....Ala 689
    : : : : : : : : : : : : : : : : : : : : : : : : : :
956 CGTCCCAATATATCGTGGCGGCTGTGTATCGCGCGCTC..... 915
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
690 ThrProGlyProAlaCysAlaSerProMetPheLysProIlePheProAl 706
    : : : : : : : : : : : : : : : : : : : : : : : : : :
914 .....AATACCGAACGGAAATCACGCGGTTGCTGTGT 881
    : : : : : : : : : : : : : : : : : : : : : : : : : :
706 aThrProLysSerGluSerAspAsnProLeupProThrSerSerSerAla 723
    : : : : : : : : : : : : : : : : : : : : : : : : : :
880 CAACCAATTCGCCGCGAGTAATTCGACTTCGCGACCCCAAAAGGTA 831
    : : : : : : : : : : : : : : : : : : : : : : : : : :
723 IatThrThrProAla.....SerThrAlaLeuProThrThrAla 736
    : : : : : : : : : : : : : : : : : : : : : : : : : :
830 CGCAAGAGCGCGGTTTGTGACTTGAGAACCCACTAGGCGCATACGCG 781
    : : : : : : : : : : : : : : : : : : : : : : : : : :
737 .....ThrAlaThrAla 740
    : : : : : : : : : : : : : : : : : : : : : : : : : :
780 CTGCGTGTTCAGAGCGCTGTTCGAAACAAAGCGCAATGGTATTCAT 731
    : : : : : : : : : : : : : : : : : : : : : : : : : :
740 aHisThrPheLysProIlePheGluSerValGluProPheAla..... 754
    : : : : : : : : : : : : : : : : : : : : : : : : : :
730 CTGTATATGATGATGTCACAGCGTTTATTCGGCGCGACCGGCTCATG 681
    : : : : : : : : : : : : : : : : : : : : : : : : : :
755 .....AlaMetProLeuSerProProPheSerLeu 764
    : : : : : : : : : : : : : : : : : : : : : : : : : :
  
```

```

680 AATGTAATGTGCTGGCACTCAAAACCGGAGATGCGCGCGCAATTC 631
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
765 LysGlnThrThrAlaProAlaThrThrAlaAlaThrSerAlaPro..... 779
    : : : : : : : : : : : : : : : : : : : : : : : : : :
630 ATGTGTTTCATGTTGGCAGACATTTTCAGACGCGCACGCTGGCCAGCTG 581
    : : : : : : : : : : : : : : : : : : : : : : : : : :
780 .....LeuLeuThrGlyLeuGlyThrAlaThrSerThrValAla 793
    : : : : : : : : : : : : : : : : : : : : : : : : : :
580 CTTTACAAACATGATTTCCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 531
    : : : : : : : : : : : : : : : : : : : : : : : : : :
793 hreGlyThrThr..... 796
    : : : : : : : : : : : : : : : : : : : : : : : : : :
530 CGCGGTTTGAATCTCGCGCGCTTCTGTGTAATGACGTAAGG..... 486
    : : : : : : : : : : : : : : : : : : : : : : : : : :
797 .....AlaSerAlaSerLysProValPheGlyPheGlyValThr 809
    : : : : : : : : : : : : : : : : : : : : : : : : : :
485 .TCGCGACCGCGGATGTGTTCATCGCATTCGATGAGAGATGGCGAAGC 437
    : : : : : : : : : : : : : : : : : : : : : : : : : :
809 rThrAlaAlaSerThrAlaSerThrThrAlaSerThrSerGlnSer.... 824
    : : : : : : : : : : : : : : : : : : : : : : : : : :
436 GCTCGGCATCGACGCGAGATTTGCTGACGCGAGG.....GTGCGC 393
    : : : : : : : : : : : : : : : : : : : : : : : : : :
825 .....IleLeupheGlyValAlaProProValThr 834
    : : : : : : : : : : : : : : : : : : : : : : : : : :
392 AGCGCATGCCAACCAACCGATTTGATCAGGTTGCGCGCGCATCTTTCGCC 343
    : : : : : : : : : : : : : : : : : : : : : : : : : :
835 AlaSerSerSerAlaProAlaLeuAlaAlaSerIlePheGlnPheGlyLysPr 851
    : : : : : : : : : : : : : : : : : : : : : : : : : :
342 GCTT...AAGTTGCGACGCGCTTCAGGTCGTAAGCTTCAAACTGCATTT 296
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
851 oleuAlaProAlaAlaSerValaIaGlyThrSerPheSerGlnSerLeuA 868
    : : : : : : : : : : : : : : : : : : : : : : : : : :
295 CGTCGTTGCTTCACGCGCATACGACGACGACGACGACGACGACGACG 246
    : : : : : : : : : : : : : : : : : : : : : : : : : :
868 IaserSerAlaGlnThrAla..... 874
    : : : : : : : : : : : : : : : : : : : : : : : : : :
245 CCACGGTGAAATCGCGCGATTTTGCCTGAAGCGCGCGCATGAACACAC 196
    : : : : : : : : : : : : : : : : : : : : : : : : : :
875 .....AlaSerAsnSer 879
    : : : : : : : : : : : : : : : : : : : : : : : : : :
195 GCCCGGATTC.....TTTGTCTTCAACACGACTGTGC 161
    : : : : : : : : : : : : : : : : : : : : : : : : : :
879 rGlyGlyPheSerGlyPheGlyGlyThrLeuThrThrSerThrSerAlaP 896
    : : : : : : : : : : : : : : : : : : : : : : : : : :
160 CTTTTCGACGCGATGCGCTCTGACTTCATTCATGAGGGGCGCATATCCG 111
    : : : : : : : : : : : : : : : : : : : : : : : : : :
896 roAlaThrThrSerGlnProThrLeuThrPheSerAsnThrValThrPro 912
    : : : : : : : : : : : : : : : : : : : : : : : : : :
110 GCATATTCCTTCGCCAAGACAGCGGACTTCGTAATGCGCGCGCGTCTGA 61
    : : : : : : : : : : : : : : : : : : : : : : : : : :
913 ThrPheAsnIleProPheSerAlaSerAlaLysProAlaLeuProThrLy 929
    : : : : : : : : : : : : : : : : : : : : : : : : : :
60 AACGCGTGTGCTCGGCTGCGCGCGCATGGGC 30
    : : : : : : : : : : : : : : : : : : : : : : : : : :
929 rProGlyAlaAsnSerGlnProThrPheGly 939
    : : : : : : : : : : : : : : : : : : : : : : : : : :
  
```

seq_name: p1r2:D83208

seq_documentation_block:
 Probable ferredoxin P43491 [imported] - Pseudomonas aeruginosa (strain PA01)
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: D83208
 R:Stover, C.K.; Pham, X.O.; Eyrln, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.;
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
 .; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
 A:Reference number: A82950; MUD:20437337
 A:Accession: D83208
 A>Status: preliminary
 A:Molecule type: DNA

A:Residues: 1-774 <STO>
 A:Cross-references: GB:AE004770; GB:AE004091; NID:g9949633; PIDN:AG06879.1; GSPDB:GN001
 A:Experimental source: strain PM01
 C:Genetics:
 A:Gene: PA3491

Alignment_scores:
 Quality: 129.00 Length: 488
 Ratio: 0.578 Gaps: 23
 Percent Similarity: 45.697 Percent Identity: 21.516

Alignment_block:

US-09-303-518D-125 x D83208 ..

Align seg 1/1 to: D83208 from: 1 to: 774

```

28 CTGCCCATCGCGGCGACACCGGACCAACCGCTTACGACGCGCGGCCCAT 77
   |||||
23 LeuProIleGlnGlnAlaProLeuAlaGlnArgTyrIleValPro..... 37
   |||||
78 TACCGAATCGCGTGTGCGCGAAGATAT...GCCGTAATGCGCCCT 124
   |||||
38 .....LeuGlnHisIleGlyAlaProAlaArgProC 49
   |||||
125 CGATGAAGTCAAGAGAGCGCATGCGTCAAAAAAGCGCAAGTGTGTT 174
   |||||
49 yValIleVal.....GlyAlaValAlaLeuLysGlyGlnThrIleAla 63
   |||||
175 GAAGCAAAAAAGATCGCGCGTGTGTACTCGCGCGCGCTTCAGCA 224
   |||||
64 LeuProAlaSerGlyThrValSerAlaAlaLeuHisAlaProThrSerGly 80
   |||||
225 AATCGCGCGCAT.....CACCGTGGCGAAAAAGCGCGTAC 259
   |||||
80 rValValAlaIleGlyAlaHisProTyrProHisIleSerGlyLeuProA 97
   |||||
260 TTCAGCATGCGTGAATGCGCTGAAGCAACGCA..... 297
   |||||
97 LeuProAlaIleValIleAlaSerAspGlyLeuGlnArgTyrThrGlnLeu 113
   |||||
298 .....ATCGAGTTTGACGCTACGACCTGAAGCGCTGCGCAACT 338
   |||||
114 HisProCysProAlaSerPheArgIleGlnSerProLeuAlaLeu..... 128
   |||||
339 AAGCGCGCAAGATCGCGCG.....AACCTGATCCATCGC 376
   |||||
129 .....GluArgIleArgAlaAlaGlyIleGlyLeuGlyGlyAlaG 143
   |||||
377 GTTGTGGACTGCG.....CTCGCACCGCGTTCAGCAAAATTCCT 420
   |||||
143 LysPheProThrAlaAlaLysLeuAlaAlaArgProAlaGlnLysIle... 158
   |||||
421 GCCGTGATGCCGACCGCTGCCCATCTTCGAATGCGATGACACCA 470
   |||||
159 .....HisThrLeuValAlaAsnGlyAlaGlnCysGln 169
   |||||
471 TCCGTCGCGTCCGACCTACGCTCATTTCAAGAAGCGCGGAGMT 520
   |||||
169 uProTyrIleSerAlaAspAspLeuMetArgGlnArgAlaThrGlnV 186
   |||||
521 TCAAACGCGCGCTGTGTTGATTCAGCGCT..... 549
   |||||
186 alleuGlyGlyIleAspIleLeuValGlnIleLeuGlyProGlnGlnVal 202
   |||||
550 .....TTGACCGAAGCAAAATTCATGTTTGAAGCGACGCGCGC 590
   |||||
203 LeuValIleGlyIleGlnAspAspLysProGlnAlaIleAlaIleLeuGlyAl 219
   |||||
591 A.....GACGTGCGCTGCAAAATGCGTCCACATCGAAGCAAG 631
   |||||
219 alaLeuGlyLeuArgProTyrArgIleValAlaLeuProThrArgTyrP 236

```

```

632 AATTCGCGCGCGCGCAT..... 648
   |||||
236 roSerGlyGlyGlnArgGlnLeuIleGlnLeuThrGlyArgGlnVal 252
   |||||
649 CTTGCC...GGTTGAGTGGCAGCAGCATTCATTCATCGACCGCGTGG 695
   |||||
253 ProAlaAspGlyLeuProAlaAspIleGlyIleLeuGlyGlnAsnValGln 269
   |||||
696 CGCGAATAAACCGTGGACCATTCATTCATTCATTCATTCATTCATTC 745
   |||||
269 yThrLeuAlaVal.....HisAspAlaValValLeuG 281
   |||||
746 GCCGTTGTTTGCACAGCGCGTGTGACACGCGCGCTGATTCGCTTA 795
   |||||
281 LysArgProLeuIleSer.....ArgIleThrThrLeu 291
   |||||
796 GGTGTTTCTCAATCAACAAACCGCGCTGTTCGTAACGTTTGGTTC 845
   |||||
292 AlaGlyGlyAlaLeuGlnArgProMetAsnValGlnAlaLeuIleGly 308
   |||||
846 GAAGTATCGCAATTCATCGC.....GGCAATTCG 877
   |||||
308 rProValHisGlyLeuLeuAlaPheAlaGlyLeuAlaGlnGlyArgLeu 324
   |||||
325 .....GluArgValLeuMetGlyGly..... 331
   |||||
928 ACACAAGCGCGCGCATATTCGGACGCTACCAATCAGATTCGCT 977
   |||||
331 ..... 331
   |||||
978 TATCGAAGAGCGCGCAGCAAGAGCTGTTCGCTGGTTCGCGCGAGC 1027
   |||||
332 .....PrometMetGlyPheAlaLeuProAsp. 340
   |||||
1028 CGGCAAAATTCATTCATTCGCGCTACACCTTCGCGCATTCCTGA 1077
   |||||
341 .....LeuSerVal 343
   |||||
1078 AAACCTTCGAATTCACACAGCGCGTCAACGCGCGCGCGCGCAATG 1127
   |||||
344 ProLeuIleLysThrCysAsnGlyLeuLeuAlaGlyAspAlaThrGln 360
   |||||
1128 GCG.....ATTCGTAATTCAGAGC 1147
   |||||
360 uProGlnProValProAlaMetProCysIleArgCysGlyAspCysAlaG 377
   |||||
1148 GCGTATGCCCTTGATATCCGCGCACCGCTGTTTCGCGGATTCATC 1197
   |||||
377 LysValCysProValSerLeuLeuProGln...GlnLeuHisPhePheAla 392
   |||||
1198 GTGCGCATACCGACAGCGCGAGCATTTGGGTTCCTGGAATTCAGCA 1247
   |||||
393 LeuGlyAsp...GlnHisGlnLeuLeuAlaHisAsnLeuPheAspCys 408
   |||||
408 sIleGlnCysGlyAlaCysAlaTyrValCysProSerSerIle..... 422
   |||||
1298 GCCGCGTGTTCGCGC 1311
   |||||
423 .....ProLeuValGln 426

```

seq_name: p1r2:A84943

seq_documentation_block:
 membrane protein [imported] - Buchnera sp. (strain Aps)
 C:Species: Buchnera sp.
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
 C:Accession: A84943
 R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
 Nature 407, 81-86, 2000

A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A
 A:Reference number: A84943; MUID:20445173
 A:Accession: A84943
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-473 <570>
 A:Cross-references: GB:AP000398; GSPDB:GN00144
 A:Experimental source: strain APS
 C:Genetics:
 A:Gene: rnlC

alignment_scores:

Quality:	126.00	Length:	449
Ratio:	0.575	Gaps:	22
Percent Similarity:	48.775	Percent Identity:	20.045

alignment_block:

US-09-303-518D-125 x A84943 ..

Align seg 1/1 to: A84943 from: 1 to: 473

```

100 GAAGAAATATGCGGGTATGCGCCCTCGATGAAATCAAGAGGCGATGC 149
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
13  AspsapTyrHisAsnValLys.....LeuArgValLysHisnGluLys 27
150 CGTCAAAAAGGCCAA...GTCTGTTTGAAGACAAAAGATCCGGCG 186
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
27  sValLeuArgLysInProLeuHisSerAspAspPheAsn.....Y 42
197 TGTGTTTACTGCGCGCTTCAGCAAAATCCGGCGATTCACCGTGC 246
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
42  alProValHisAlaProHisSerLysLeuLysLysLysLysLysLys 58
247 GAAAGGCGGCTACTTCAGTCAAGTGTGATTCGGCTTGAGGC..... 288
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
59  SerAspSerLysLysAsnLysLysLysLysLysLysLysLysLys 75
289 .....AACGACGAAATCGAGTTTGAAGCGTACGCAC 319
75  uAspGlnTTrpLysArgLysAsnProLysAspTyrLysLysLysLys 92
320 CTGAAGGCGTGGCAAACTTAAGCGCGAAGAGTGGCGGCAACCTGATC 369
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
92  roGluLysLysLysLysLysLysLysLysLysLysLysLysLys 100
370 CAATCCGGTGTGGACTGCGTGGCGACCCGCTGCCTC.....AG 410
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
101  GlnSerGlyVal...ValGlyLeuGlyGlyGlyGlyGlyGlyGly 116
411 CAAATTCCTGCGCGATGCGAGCGGCTTGGCATCTCGTCAATGAGA 460
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
116  sLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 133
461 TGGACACCAATCGCTGCGTGGCGACCGCTACGCTATTATCAAGAAGCC 510
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
133  alGlnSerGlnProLysLysLysLysLysLysLysLysLysLys 149
511 GCCGAGGATTTCAACGCGCGCTGTGATTTGAGCGGCTTACCGAAGC 560
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
150  HisSerGlnLysLysLysLysLysLysLysLysLysLysLysLys 166
561 CAAATTCATGTTTGAAGGACGTCGGCGACGCTGCGCTCT..... 603
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
166  elYsThrValLeuLysLysLysLysLysLysLysLysLysLysLys 183
604 .....GAAATGCTGCCACATGGAACACATGAATTC 636
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
183  ystLeuHisLysLysLysLysLysLysLysLysLysLysLysLys 199
637 GCGGCGCGCATCTGCGCGT.....TTGAG 662
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
200  LysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 216
  
```

```

663  TGGCAGCCACATT.....CATTTATCAGAGCCGTCGCGCA 700
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
216  rGlyLysGluValProHisLysLysHisSerLysLysLysLysLys 230
701  ATAAACCGTGTGACCATTAATTCAGATTAATTAATTCAGATTAATTC 750
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
231  .....TyrLeuLysLysLysLysLysLysLysLysLysLysLys 244
751  TTGTTTGGACACAGCGCGCTGAAACAGCGCGCTGATTCGCTAGTGG 800
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
245  AlaLysLysLysLysLysLysLysLysLysLysLysLysLysLys 259
801  TTCTCAAGTCAACAAACCGCGCTGTCGCTACCGCTTTGGTGGCAAG 850
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
260  .....MetSerAspLysAsnLeuLysLysLysLysLysLysLys 274
851  TATCGCAA.....ATTACTGCGCGGCAATTTGTTGACACA 885
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
274  LeuLysThrProLysLysLysLysLysLysLysLysLysLysLys 290
886  GACACCGCGTATTTCCGGTTCGTTATTAACGCGCGCTATTCACAAAG 935
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
291  .....PheLeuLysLysLysLysLysLysLysLysLysLysLys 299
936  CGGCGACGATTTTGGGCGCTACCAACATGATTTCCGTTATTCAGAG 985
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
300  .....ProPheMetGlyLysLysLysLysLysLysLysLysLys 310
986  AAGGCGCGACGAAAGAGTTCGCGCTGCGTTCGCGCGCGCGACGCA 1035
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
310  ..... 310
1036  TACTGCATCAGCGGTACACCGCTGCGCATTTCTGAAAACAACTCTT 1085
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
311  HisSerLysLysLysLysLysLysLysLysLysLysLysLysLys 327
1086  CAAATTCACACAGCGCTCAACAGCGCGCGCGCGCGCGCGCGCGAG 1135
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
327  uLys.....AsnGlnSerLysLysLysLysLysLysLysLysLys 341
1136  GTACTTACGAGCGCGTATGCGCTTGCATATCTGCCACCGCTGCT... 1182
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
341  LysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 357
1183  .....TGGCGGATTTAATCGCGCGATTCGACGACGCGCGCGCAT 1226
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
358  TrpTyrLysLysLysLysLysLysLysLysLysLysLysLysLys 374
1227  GGGTTCCTGGAATTTGACGAAAGACCTTCCTTGTGACGCTTCGCT 1276
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
374  uAspCysLysLysLysLysLysLysLysLysLysLysLysLysLys 386
1277  GCCCGGGCAATACGATACGCGCGCGCTGTTGGCAAGAGCTGAGAA 1323
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
386  ySPro.....SerTyrLysLysLysLysLysLysLysLysLysLys 398
seq_name: p1r2:T39903
seq_documentation_block:
serine-rich protein - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Nov-2000
R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Rieger, M.
submitted to the EMBL Data Library, November 1998
A:Reference number: Z21889
A:Accession: T39903
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-534 <LYN>
A:Cross-references: EMBL:AL033534; PIDN:CAA22127.1; GSPDB:GN00067; SPDB:SPBC215.13
A:Experimental source: strain 972h; cosmid c215
  
```


Align seg 1/1 to: S01359 from: 1 to: 328

```

29  TCCCATCGCGGCGAGCGAGCAAGCGTTTACGAGCGCCG... 73
   |||||:|||||:|||||:|||||:|||||:|||||:
28  CysProLysArgThr...ThrProLysProCysThrThrAlaAlaProTh 43
   .....CCATTACCGAAGTCGCGCTTGGCGAATAATGCGGAT 116
   |||||:|||||:|||||:|||||:|||||:|||||:
43  rCysAlaProValThrThrThrCysArgProProThrThrThrArgC 60
   |||||:|||||:|||||:|||||:|||||:|||||:
117  GCGCCCTCGATGAAGTCAAGGAAGCGATCGTCAAAAAAGCCAG 166
   |||||:|||||:|||||:|||||:|||||:|||||:
60  ySPioProThrThrThrArgCysProProProThrArgProAlaGlu 76
   |||||:|||||:|||||:|||||:|||||:|||||:
167  TCGCTTTGAAGACAAAAGATCCGCGCTGTGTACTGCGCCG... 214
   |||||:|||||:|||||:|||||:|||||:|||||:
77  CysThrAlaThrThrLysArgProThrAlaArgProThrThrAlaGlyTh 93
   .....CTTACGCAAAATGCGCGATTCACCGTGGCGAAGCGCGTACT 260
   |||||:|||||:|||||:|||||:|||||:|||||:
215  ...CTTACGCAAAATGCGCGATTCACCGTGGCGAAGCGCGTACT 260
   |||||:|||||:|||||:|||||:|||||:|||||:
93  rThrValArgAlaThrThrLysArgAlaThrThrArgThrThrLysA 110
   |||||:|||||:|||||:|||||:|||||:|||||:
261  TCAGTCAGTCG...TGATTGCGCTTGAAGCAAGCAAGCAAA...TCGAGT 304
   |||||:|||||:|||||:|||||:|||||:|||||:
110  rGAlaThrThrArgArgThrThrValArgAlaThrThrLysArgAlaThr 126
   .....TTGAAGCTACGACCTGAAAGCGCTGGCAACT...TAAGCGCGAGAA 351
   |||||:|||||:|||||:|||||:|||||:|||||:
127  ThrArgArgThrThrThrLysArgAlaProThrArgThrThrThrLys 143
   |||||:|||||:|||||:|||||:|||||:|||||:
352  GTGGCGCGCAACTGATCAATCCGCTT...TGAGACTGCGTGGCG 395
   |||||:|||||:|||||:|||||:|||||:|||||:
143  sArgAlaThrThrArgArgAsnProThrArgArgThrThrThrArgArg 160
   |||||:|||||:|||||:|||||:|||||:|||||:
396  CACCGTCGCTTCAAGAAATTCCTGCGCGTATGCGAGCGCTGCCCA 445
   |||||:|||||:|||||:|||||:|||||:|||||:
160  lAProThrLysArgAlaThrThrLysArgAlaThrThrArgArgAsnPro 176
   .....TCGTCATGCGATGAGCAACCAATCCG... 475
   |||||:|||||:|||||:|||||:|||||:|||||:
177  ThrLysArgLysThrThrArgArgThrThrValArgAlaThrLysThrTh 193
   .....TGCGTCGCGACCGCTACGCTCATTTACAAGA 506
   |||||:|||||:|||||:|||||:|||||:|||||:
476  .....TGCGTCGCGACCGCTACGCTCATTTACAAGA 506
   |||||:|||||:|||||:|||||:|||||:|||||:
193  rLysArgAlaThrThrLysArgAlaProThrLysArgAlaThrThrLysA 210
   .....AGGATTTCAACCGCGCTGTGTGTAATGAGCGCT 550
   |||||:|||||:|||||:|||||:|||||:|||||:
507  AGCGCGCG...AGGATTTCAACCGCGCTGTGTGTAATGAGCGCT 550
   |||||:|||||:|||||:|||||:|||||:|||||:
210  rGAlaProThrLysArgValThrThrLysArgAla... 221
   |||||:|||||:|||||:|||||:|||||:|||||:
551  TGACCGAAGCAAAATCCATGTTGTAAAGCAGCTGCGCAGACGTCGCG 600
   |||||:|||||:|||||:|||||:|||||:|||||:
222  ...ProThrLysArgAlaThrThrLysArgAlaProThrLysArgAlaTh 237
   |||||:|||||:|||||:|||||:|||||:|||||:
601  TCTGAATACTGCGCACATCGAACAACATGATGCGCGCGCCGATCC 650
   |||||:|||||:|||||:|||||:|||||:|||||:
237  rThrLysArgAlaProThrLysArgAlaThrThrLysArgAlaProThrL 254
   .....TGCGGCTTGTAGTGCGACGC... 670
   |||||:|||||:|||||:|||||:|||||:|||||:
651  TGCGGCTTGTAGTGCGACGC... 670
   |||||:|||||:|||||:|||||:|||||:|||||:
254  ysaArgAlaThrThrLysArgAlaProThrLysArgAlaThrThrLysArg 270
   .....ACATTCATTCATCGACCGCGCGCGGATTAACCGTGTGAGCC 717
   |||||:|||||:|||||:|||||:|||||:|||||:
271  AlathrAlaArgProThrSerLysProCysGlyCysLysProCysGlyPr 287
   |||||:|||||:|||||:|||||:|||||:|||||:
718  A 718
287  0 287
seq_name: p1r2:S59310

```

seq documentation block:
 Problem membrane protein YMR317w - yeast (*Saccharomyces cerevisiae*)
 N:Alternate names: hypothetical protein Y9924.09
 C:Species: *Saccharomyces cerevisiae*
 C:Date: 29-Nov-1995 #sequence_revision 23-Feb-1996 #text_change 05-Nov-1999
 C:Accession: S59310
 R:Churcher, C.M.
 Submitted to the EMBL Data Library, September 1995
 A:Reference number: S59302
 A:Accession: S59310
 A:Molecule type: DNA
 A:Residues: 1-1104 <CHU>
 A:Cross-references: EMBL:Z54141; GSPDB:GN00013; MIPS:YMR317w
 A:Experimental source: strain AB972
 C:Genetics:
 A:Gene: MIPS:YMR317w
 A:Map position: 13R

alignment_scores:
 Quality: 124.00 Length: 466
 Ratio: 0.588 Gaps: 21
 Percent Similarity: 45.279 Percent Identity: 23.176

alignment_block:
 US-09-303-518D-125/rev x S59310 ..

Align seg 1/1 to: S59310 from: 1 to: 1104

```

1331 TCATAGTGTTCACACCTTTGCGCAACAGCGGCGGTATTCGATTATTCGC 1282
   |||||:|||||:|||||:|||||:|||||:|||||:
350  SerSerThrSerSerSerValSerSerGluAlaProSerSer... 363
   |||||:|||||:|||||:|||||:|||||:|||||:
1281  CGGGCAGACGAGCTGCACAAAGCCAGCTCTCTTCGTCATTCGACAG 1232
   |||||:|||||:|||||:|||||:|||||:|||||:
364  .....ThrSerSerSerValSerSerGluA 372
   |||||:|||||:|||||:|||||:|||||:|||||:
1231  AACCCATGCTGCGCGCTGCTGTATGCGCGACGATTAAATCGCGCAAA 1182
   |||||:|||||:|||||:|||||:|||||:|||||:
372  lAProSerSerThrSerSerSerValSerSerGluAlaSerSerThrLys 388
   |||||:|||||:|||||:|||||:|||||:|||||:
1181  AAGCAGGTGGCAGCATATTCACAGGCGATCACGCCCTCGTAAGTACCAAT 1132
   |||||:|||||:|||||:|||||:|||||:|||||:
389  SerSerVal...MetSerSerGluValSerSerAla...Th 400
   |||||:|||||:|||||:|||||:|||||:|||||:
1131  CGGACCATGCGCGCGCGCGCGCGCTTGCAGCGCTGTGAATGAGTA 1082
   |||||:|||||:|||||:|||||:|||||:|||||:
400  rSerSerLeuValSerSerGluAlaProSerAlaAlaSerSerLeuAla 417
   |||||:|||||:|||||:|||||:|||||:|||||:
1081  GT...TTGTTTTCAGAAATGCGCGAGGCTGTACGCGTGTG... 1041
   |||||:|||||:|||||:|||||:|||||:|||||:
417  rSerArgLeuPheSerSerLysAsnThrSerValThrSerThrLeuVal 433
   |||||:|||||:|||||:|||||:|||||:|||||:
1040  .....GAGTATTTTCGCGCTGCGCGCGCAACCCAGCGCAACAGCTCTT 997
   |||||:|||||:|||||:|||||:|||||:|||||:
434  AlathrGluAlaSerSerValThrSerSerLeuArgProSerSerGluTh 450
   .....GCTGCGCGCTTCTGATACGGAATGATGATTGTGTAGGCTCCCAAT 947
   |||||:|||||:|||||:|||||:|||||:|||||:
946  rLeuAlaSerSerSerLeuIleGlu... 458
   |||||:|||||:|||||:|||||:|||||:|||||:
946  NATCGTGGCGCGCTTGTGTAATCGCGCGCTTCAATACCGAACCGAATATC 897
   .....ACGCGCTTGTCTGTACCAACATTCGCCCGCAGTAAATTCGATCTTT 847
   |||||:|||||:|||||:|||||:|||||:|||||:
846  CGCACCAAAAGGTACGACGAGCGCGGCTTTGTGACTGGAACACCAC 797
   |||||:|||||:|||||:|||||:|||||:|||||:
475  aAlaSerSerThrLeuGlySerLys... 483

```

```

796 CTAGGGCAATCAACGGCTGGTTCAGACGGCTTTGCAACAACAGCG 747
484 .....ValSerSerSerSerSerArgMetalThrSerThr 496
746 CCNAATGTAATTAACATCTGTAATTAATGATGTCACACGGTTTATGCG 697
497 SerSer.....ThrSerSerSerSerSerSerSerSerValIlePhe1 511
696 GCCGACGGCTCGATGAATGATGCTGCTGCACTCAACACGGCAGAT 647
511 YasnSerSerThrVal.....T 517
646 GCGGCGCGCGAATTCATGTTTGATGATGTCAGCA..... 609
517 hrThrSerProSerAlaSerIleSerLeuThrAlaSerProLeuProSer 533
608 .....TTTTCAGACGGCAGCTGTGCGCCAGCTGCTTACAAACATGATTTT 562
534 ValTrpSerAspIleThrSerSerGluAlaSerSerIleSerSerAsn1 550
561 GCGTTCGGTCAACGGCTCAATACACAGACGGCGCTTGAATCTCTCG 512
550 uAlaSerSerSerAlaProSerSerAspAsnSerThrIleAlaSer.... 565
511 GCGCTCTTGAATGATGAC.....GTAGGCTGCGCA 480
566 ..AlaSerLeuIleValThrLysThrLysAsnSerValValSerIle 581
479 GCGACGGATGGTGTTCATGCAATGACGATGACGAGATGGCGACGTCGCG 430
582 ValSerSerIleThrSerSerGluThrAsnGluSerAsnLeuAlaThr 598
429 ATCGACGGCAGAAATTTGCTGAACGACGCGGTGCGCAGCGTCACAC 380
598 rSerSerThrSerLeuSerAsn..... 606
379 AACCGATTGATCAGTTCGGCGACCTTCGCGGCTTAAGTTTGGC 330
607 .....LysAlaThrAlaArgSerLeuSerThrSer 616
329 AGCGCTTCAGGTGCGTACGCTTCAAC.....TCGATTTGCTC 292
617 AsnAlaThrSerAla.....SerAsnValProThrGlyThrPheSer 631
291 GTTGCCTTCA...ACGGCAATCAAGACTGACTAAGTACCGCTTTTCCG 245
631 rMetSerSerThrSerVal..... 638
244 CACGGTAATCGCGGATTTTCCCTGAAGCGCGCAGCAATAACACACG 195
639 .....IleThr 640
194 CCGGATTCCTTTTGTCTCA.....AACGCACTTGGCC 160
641 ProGlyPheSerThrSerSerAlaSerLeuAlaIleAsnSerThr.... 655
159 TTTTTCAGCGGATCGCTTCCTGACTTCATGAGAGGGGCGATACCG 110
656 .....ValValSerSerSerLeuAlaGlyTyrSerPheSerThrPro 670
109 CATATCTTCGCCAAGCAACGCACTTCGTAATGCGCGGCGCTGTAA 60
670 Lu...SerSerProThrThrSerThrLeuValThrSer..... 681
59 ACGGCTTCCTCCGCTGCGCGCAGATGGAGGTTTAGACTTTTGG 12
682 GluAlaProSerThrValSerSerMetThrThrSerAlaProPheIle 697
seq_name: plir2:T05352
seq_documentation_block:
hypothetical protein FB84.120 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)

```

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999
C:Accession: T05352
R:Bevan, M.; Terry, N.; Ardiles, W.; Buyschaert, C.; Dasseville, R.; De Clerck, R.;
ewes, H.W.; Mayer, K.F.X.; Schueller, C.
Submitted to the Protein Sequence Database, February 1999
A:Reference number: Z15409
A:Accession: T05352
A:Molecule type: DNA
A:Residues: 1-857 <BEV>
A:Cross-references: EMBL:AL034567
A:Experimental source: cultivar Columbia; BAC clone FB84
C:Genetics:
A:Map position: 4
A:Insertions: 26/3; 45/1; 74/3; 83/1; 122/2; 165/1; 270/2; 307/1; 731/2; 754/2
A:Note: FB84.120
C:Superfamily: cyclophilin homology
F:6-162/Domain: cyclophilin homology <CYP>

alignment_scores:
Quality: 118.00 Length: 393
Ratio: 0.663 Gaps: 22
Percent Similarity: 45.293 Percent Identity: 25.700

alignment_block:
US-09-303-518D-125 x T05352 ..

Align seg 1/1 to: T05352 from: 1 to: 857

```

24 AAACCTGCCATCGCGGCGACGACGAGCAAGCCGTTTACAGCGGCCGG 73
520 ArgSerAlaSerArgGlySerLeuGlyArgGlyProLeuArgArgSer 536
74 CCATTACGGAAGTCCGTTGCTTGGCGAAGATATGCGGTATGCGGCC 123
536 r.....ArgArgSerProSerArgSerPro 545
124 TCGATGAATGCAAGAGCGCATCGCTCAAAAAGGCAAGTCTGTT 173
545 alArgSerSer.....ArgArgSerLeuSerArgSerProIleGlnLeu 559
174 TGAAGCAAAAAGAAATCCGGCGCTGTCTTCTCGCGCGGCTTACGCA 223
560 .....SerArgArgSerLeuSerArgSerProThrArgLe 571
224 AAATCGCCCGCAT.....TCACCGTGGCGAAAGCGCTACTGAG 264
571 uSerArgArgSerLeuSerArgSerProIleArgSerProArgGlySer 588
265 TCAGTGTGATGTCGCTTGAAGGCAACGCAAAATGAGTTGAACGCTA 314
588 alSerArgSer..ProValArgSerSerArgGlySer..... 599
315 CGACCTGAAGCGCTGCAAACTTAAGCGGCGAAGAGTGGCGCAACC 364
599 ..... 599
365 TGATCCAAATCGGTTTGTGACTGCGCTCGACCGCTGCTTCAACA.. 412
600 .....ValSerArgSerProValArgSerSer 609
413 ..AAATTCGCGCGTGCATGCCGAGCCGTTGCCATCTTGTCAATGCCAT 461
609 gArgArgIleSerArgSerProValArgSerSerArgGlySerValSera 626
462 GGAACCAACATCCGCTGCTGCGACGACCTTCGATTAACAAGAGCCG 511
626 rSerProIleArgLeu.....SerArgArgSerIleSerArgSerPro 640
512 CCGAGATTTCAACGCGCGCTTGTGATGACCG..... 548
641 IleArgLeuSerArg..ArgSerIleSerArgSerProValArgGlyArg 657

```

```

549 .....TTGACCGACGCAAAATTCATGTTGTAGGACGCTGGCCAGA 593
557 rglrllserargserprovalproala.....Argarg 668
594 CGTGGCTGTGAATAATGCTGCCAATCGAANACATGATTCGGGGGCC 643
669 ArgSerValArgPro.....ArgserPr 676
644 CGCATCCCTGGCTTGTAGTGACGACGACATTCATTCGAGCGCGTC 693
676 OProProAspArgArgArgSerLeuSerAlaSerProAsnGlyA 693
694 GCGCGCAATAAAACGCTGTGACCATCATTCATCAAGATGATTAACAT 743
693 rglleargargllyargllypsergin.....Argpserlyr... 706
744 TGGCCGTTGTTGCAACGCGCTGCAACGCGCGCGCTGATTCGCC 793
707 .....AlaargarglyrarghSerPr 714
794 TAGGTGTTCTCAAGTCACAAACCGCGCTTCTGCGTATTCGGGT 843
714 o.....SerProAspArgSerPro.....Tyrargpser. 724
844 GCGAAGTATCGCAATTAATCTCGGGCGATTTGTTGACACAGCAACCG 893
725 .....AspArgSerasp 728
894 CGTGAATTCGGCTGATTTGACGCGCGATTCACACAGCGCGCAGC 943
729 Argasp.....ArgpserArgargpserProserArgpser 744
944 A.....TTATTTGGACGCTACCAATCATCATTCCTGATGAGAA 987
744 gSerProleuargllyarghProProSerMetleuargllyargh 761
988 GCGCGCACAAAGAGCTGTGCGGTGCGCGCGCGCGCAATA 1037
761 rg.....SerArgSerValSerProgllyLeuCYTyrArgAsnArg... 774
1038 CTCATCAGCGCTACACCT.....CGGCCATTTCTCTGA 1072
775 .....ArgTyrSerArgSerProIleArgSerArgSerProIlyrargly 790
1073 AAACAAACTCTCAAGTTCACACAGCGCGCGCGCGCGCC 1122
790 sarArgSerProSerAlaSerLeuSerProSerArgSerArg 807
1123 ATGTCGCCGATTTGTTGACTTACGAGC 1148
807 erArgSerlySerlySerlySer 815
seq_name: p1r2:G64919

```

seq_documentation_block:
probable iron-sulfur protein b1629 - Escherichia coli
M:Alternate names: nrfc protein homolog b1629
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 17-Mar-2000
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; CC
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: G64919
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residue type: 1-740 <BLAT>
A:Cross-references: GB:AE000258; GB:U00096; NID:g2267121; PIDN:AACT4701.1; PID:g1787916;
C:Experimental source: strain K-12, substrain W6165
C:Superfamily: unassigned ferredoxin 2[4Fe-4S]-related proteins; ferredoxin 2[4Fe-4S] hc
C:Keywords: iron-sulfur protein

```

F:370-434/Domain: ferredoxin 2[4Fe-4S] homology <FER>
Alignment_scores:
  Quality: 115.00      Length: 322
  Ratio: 0.728        Gaps: 12
  Percent Similarity: 49.068      Percent Identity: 22.050
alignment_block:
  US-09-303-518D-125 x G64919
Align seg 1/1 to: G64919 from: 1 to: 740
100 GAAGATATATGCGGTATGCGCCCTCGATCAATGAAGTCAAGAGCGCATC 149
48 LysInHisIleGlyAlaGluGlyGluLeuGlyValSerValIleAspLys 64
150 CGTCAAAAAGAGCCCAAGTCTGTTGAGACAAAGAAAGATCCGGCGTGG 199
64 SValleuArgGlyGlnProleuThrArgGlyArgGlyLysMetLeuPro 81
200 TGTTTACTGCGCGCGCTTCAGCAAAATCGCGCGAT..... 237
81 alHis...AlaProThrSerGlyThrValThrAlaIleAlaProHisSer 96
238 .....CACCGTGCGCAAAAGCGCGTACTTCAGTCACTGCTGATTCGCT 281
97 ThrAlaHisProSerAlaLeuAlaGluLeu...SerValIleIleAspAl 112
282 TGAGGCGACGAC..... 294
112 aaSpGlyGluAspCysTrpIleProArgAspGlyTrpAlaAspTyrArg 129
295 .....GAATCGAGTTTGACGCTAGCGCTACCGACTCAACCGCTGCA 333
129 hrArgSerArgGluGluLeuIleGluArgIleHisGlnPheGlyValAla 145
334 AACTTAAGCGGCGGAGAGAGTGGCGCGCAACTGATCAATCCGGTTGTG 383
146 GlyLeuGlyGly.....AlaGlyPhePr 153
384 GACTGCGCTGCGCACCGCGCTTCAGCAAAATTCGCGCGTGCATGCGC 433
153 oThrGlyVal.....LysLeuGlnGlyGlyAsp 164
434 AGCGGTTCGCATCTTCGTCATTCGATGAGAGACCAACATCCGTGCTGCC 483
164 yslleGluThrleuIleleasnaAlaIleGluCysGluProTyrIleThr 180
484 GACCTTACGCTATTCATCAAGAGAGCGCGCGATTCMAAGCGCGGCT 533
181 AlaAspAspArgLeuMetGlnAspCysAlaIleAlaIleValGluGlyI 197
534 GTTGTATTTGACCGCTTTGACCGAA..... 558
197 eArgIleleuAlaHisIleleuGlnProArgGluIleleuIleGlyIle 214
559 .....CGCAAAATTCATGTTGTAAGCGCGCTGCGCGCAGC 594
214 LuAspAsnLysProGlnAlaIleSerMetLeuArgAlaValAlaLeuAsp 230
595 GTGCCGTCTGAATAATGCTGCGCAACATCGAACAATCATTCGCGCGCC 644
231 .....SerAsnAspIleSerleuArgValIleProThrly 242
645 GCATCTCCCGCT.....TTGAGTGGCAGCG 670
242 styPrProSerGlyGlyAlaLysGlnleuThrTyrIleleuThrGlyLys 259
671 ACATTCATTCATGAGCGCGGTGCGCGCAATAAACCGTGTGACCATC 720
259 Inval.....ProHisGlyLysArgSerSerAspIleGlyVal 271

```


A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-772 <HAV>
 A:Cross-references: GB:BA000007; PIDN:BA035761.1; PID:q13361805; GSPDB:GND0154
 A:Experimental source: strain 0157:H7, substrain RMD 0509952
 C:Genetics:
 A:Gene: ECG32338
 C:Superfamily: unassigned ferredoxin 2[4Fe-4S]-related proteins; ferredoxin 2[4Fe-4S] h

Alignment_scores:
 Quality: 115.00 Length: 322
 Ratio: 0.728 Gaps: 12
 Percent Similarity: 49.068 Percent Identity: 22.050

Alignment_block:
 US-09-303-518D-125 x B90921 ..

Align seg 1/1 to: B90921 from: 1 to: 772

```

100 GAAGAAATATGCGCGTATGCGCCCTGATGAAAGTCAAGAGAGCGATGC 149
    : : : : : : : : : : : : : : : : : : : : : : : : : :
48 LysGlnHisIleGlyAlaGlyGluGlyLeuGlyValSerValGlyAspGly 64
    ||| : : : : : : : : : : : : : : : : : : : : : :
150 CGTCAAAAAGAGCGAGTGTGTTGAAGACAAAAGATCCGGCGCTGG 199
    ||| : : : : : : : : : : : : : : : : : : : : : :
64 sValLeuArgGlyGlnProLeuThrArgGlyArgGlyLysMetLeuProV 81
    ||| : : : : : : : : : : : : : : : : : : : : : :
200 TGTCTACTGCGCGCTTCAAGCAAAATCCCGCGATT..... 237
    ||| : : : : : : : : : : : : : : : : : : : : : :
81 AlHis...AlaProThrSerGlyThrValThrAlaIleAlaProHisSer 96
    ||| : : : : : : : : : : : : : : : : : : : : : :
238 .....CACGTGGCGAAAGCCGCGTCTTCACTGCTGATGTCGCGT 281
    ||| : : : : : : : : : : : : : : : : : : : : : :
97 ThrAlaHisProSerAlaLeuAlaGluLeu...SerValIleLeuAspAl 112
    ||| : : : : : : : : : : : : : : : : : : : : : :
282 TGAAGGCAAGAC..... 294
    ||| : : : : : : : : : : : : : : : : : : : : : :
112 aaApGlyGluAspGlyTrpIleProArgAspGlyTrpAlaAspTrpArgS 129
    ||| : : : : : : : : : : : : : : : : : : : : : :
295 .....GAATCGAGTTTGAACGCTACGACCTGACGCGCTGCGA 333
    ||| : : : : : : : : : : : : : : : : : : : : : :
129 eArgArgArgGlyGluGlyLeuIleGlyAlaGlyIleHisGlnPheGlyAla 145
    ||| : : : : : : : : : : : : : : : : : : : : : :
334 AACTTAAGCGCGGAGAGAGTGGCCGACACCTGATCCAAATCCGGTTTGG 383
    ||| : : : : : : : : : : : : : : : : : : : : : :
146 GlyLeuGlyGly.....AlaGlyPhePr 153
    ||| : : : : : : : : : : : : : : : : : : : : : :
384 GACTGGCGTGGCGACCCGCTCCGTTCAAGAAATTCCTGCGCGATGCCG 433
    ||| : : : : : : : : : : : : : : : : : : : : : :
153 oThrGlyVal.....LysLeuGlnGlyGlyAspL 164
    ||| : : : : : : : : : : : : : : : : : : : : : :
434 AGCGGTTCCCATCTTCGTAATGCGAGTGCACCAATCCGCTGGCTGCC 483
    ||| : : : : : : : : : : : : : : : : : : : : : :
164 yslIleGlyThrLeuIleLeuAsnAlaGlyGlyGlnProGlyTrpIleThr 180
    ||| : : : : : : : : : : : : : : : : : : : : : :
484 GACCCATCGCTATTCATCAAGAGCGCGGAGTTTCAACGGCGGCT 533
    ||| : : : : : : : : : : : : : : : : : : : : : :
181 AlaAspArgArgLeuMetClnAspGlyAlaGlnAlaValGluGlyIle 197
    ||| : : : : : : : : : : : : : : : : : : : : : :
534 GTTGGATTGAGCCGTTTGACCGAA..... 558
    ||| : : : : : : : : : : : : : : : : : : : : : :
197 eArgGlyLeuAlaHisIleLeuGlnProArgGlyIleLeuIleGlyIleG 214
    ||| : : : : : : : : : : : : : : : : : : : : : :
559 .....CGCAAAATCCATGTTTGAAGCGAGCTGGCGGACAG 594
    ||| : : : : : : : : : : : : : : : : : : : : : :
214 LuAspAsnLysProGlnAlaIleSerMetLeuArgAlaValLeuAlaAsp 230
    ||| : : : : : : : : : : : : : : : : : : : : : :
595 GTGCGGTGTGAAGATGCTGCACATCGAAGACATGTAATTCGGCGGCC 644
    ||| : : : : : : : : : : : : : : : : : : : : : :
231 .....SerHisAspIleSerMetArgValIleProHnL 242
    ||| : : : : : : : : : : : : : : : : : : : : : :
645 GCATCCTCGCGGT.....TTGAGTGGCAAGC 670
    ||| : : : : : : : : : : : : : : : : : : : : : :

```

seq_name: p1r2:T43481

seq_documentation_block:

Probable mucin DKFZp434C196.1 - human (fragment)

N:Alternate names: protein DKFZp434B0635.1

C:Species: Homo sapiens (man)

C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 20-Apr-2000

C:Accession: T43481; T34549; T17264

R:Koehrer, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, December 1999

A:Reference number: Z22514

A:Accession: T43481

A:Molecule type: mRNA

A:Residues: 1-580 <AAAs>

A:Cross-references: EMBL:AL133561; NID:g6599133; PIDN:CAB63715.1; PID:g6599134

A:Experimental source: adult testis; clone DKFZp434C196

R:Poustka, A.; Wellenreuther, R.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, October 1999

A:Reference number: Z21540

A:Accession: T34549

A:Molecule type: mRNA

A:Residues: 262-580 <POU1>

A:Cross-references: EMBL:AL122069; NID:g6102864; PIDN:CAB59245.2; PID:g7018420

A:Experimental source: adult testis; clone DKFZp434B0635

R:Poustka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, September 1999

A:Reference number: Z18723

A:Accession: T17264

A:Molecule type: mRNA

A:Residues: 262-580 <POU2>

A:Cross-references: EMBL:AL117481; NID:g5911958; PIDN:CAB55954.1; PID:g5911959

A:Experimental source: adult testis; clone DKFZp434B061

C:Genetics:

A:Note: DKFZp434C196.1; DKFZp434B0635.1; DKFZp434B061.1

Alignment_scores:

Quality: 114.50 Length: 563
 Ratio: 0.540 Gaps: 25
 Percent Similarity: 37.635 Percent Identity: 21.492

Alignment_block:

US-09-303-518D-125/rev x T43481 ..

Align seg 1/1 to: T43481 from: 1 to: 580

```

1336 CCTTCTCAATGGTTTCACGACCTTTCGGCAACGCGGCCGCTATTCGTAT 1287
1337 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
29 ProSerArgAlaSerProThrArgMetProSerArgAlaSerLeuLysMet 45
1286 TTGGCCGGGCGAGAGAGCTGCACAAAG.....CGAGCTCTCTTC 1246
1287 : ||| ||| ||| ||||| : ||||| ||||| ||||| |||||
45 ThrProheArgAlaSerLeuThrLysMetGluSerThrAlaLeuLeu 62
1245 GTCCAAATTCGA..... 1235
1246 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
62 rGThleuProArgAlaSerLeuMetArgThrProThrArgAlaSerLeu 78
1234 .....AG.1233
79 MetArgThrProProArgAlaSerProThrArgLysProProArgAlaSe 95
1232 CAACCCATGGCTGGCGCTGCGTATCCGCCAGCATTAATACGCGCA. 1184
1233 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
95 rProArgThrProSerArgAlaSerProThrArgArgLeuProArgAlas 112
1183 .....AAA 1181
112 ePromeGlySerProHisArgAlaSerProMetArgThrProProArg 128
1180 GCAGGCTGGCGAGATATCCAG.....GGCATCAGCGC 1147
1181 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
129 AlaSerProThrGlyThrProSerThrAlaSerProThrGlyThrProse 145
1146 CTCGTAAGTCAATCGCAGCATGCGCGG.....TCGC 1112
1147 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
145 rSerAlaSerProThrGlyThrProProArgAlaSerProThrGlyThrP 162
1111 CGCGCTGACGCGCTGTGAAGTGAAGAGTTGTTTTCAGCAAAATGG 1062
1112 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
162 rPro.....ArgAlaThr 166
1061 CGAGGCTGTACGCGTATGAGATATTTCGCCGCTGCGCGCAACCCA 1012
1062 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
167 .....AlaThrAr 169
1011 GCCGAAACG.....TCTTTCGCGCGCTTCCTTCGATAACGGAATCT 968
1012 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
169 gSerProSerThrAlaSerLeuThrArgThrProSerArgAlaSerLeuT 186
967 GATTGCGTAGCGTCCCAATATGCTGCGCGCTGTGTATATCGCGCGC 918
968 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
186 hrArgThr.....ProProArgAlaSerPro 194
917 TTCATACCGAACCG...GAAATCAGCGCGTTGTCT.....GTGTC 880
918 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
195 ThrArgThrProProArgLysSerProArgMetSerHisArgAlaSerPr 211
879 AACCAATGCGCGCAGTAAATTCGATCTTCGCACCCAAACGGTAC 830
880 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
211 ofhrArgThrPro.....ProArgAlaSerProThrArgA 223
829 GCAGAGGCGCGCTTGTGACTTGAGAACCACTGACATTCACGCGC 780
830 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
223 rGProProArgAlaSerProThrArgThrProProArgLysSerLeuArg 239
779 TCGGTGTTTCA.....CGGCTGTTGCAA 754
780 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
240 ThrSerHisArgAlaSerProThrArgMetProProArgAlaSerProth 256
753 CAAGGCGCAATGTAATTCATCTTGATTAATGATGTCACACGGTTT 704
754 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
256 rArgArgPro.....ProA 261
703 TATTGCGCGAGCGCTCGATGAATGAATGCGTCCGACCTCAACCG 654
704 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
261 rGAlaSerProThrGlySerProPro...ArgAlaSerProMetThrPro 276
653 GCAGGATGCGGG.....CC 640

```

```

277 ProArgAlaSerProArgThrProProArgAlaSerProThrThrThrPr 293
278 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
639 GCCGAATTCATGTTTCGATGTTGCACGACATTTTCAGACGCGACGTCG 590
640 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
293 oSerArgAlaSerLeuThrArgThrProSerThrAlaSerProThrThrT 310
589 CGCCAGCTGCTTCAACATGATTTTGGTTCGGTCAACAGGCTCAAT 540
590 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
310 hrProSerArgAlaSerLeuMetLysMetGluSerThrAlaSerLeuTr 326
539 ACCAAGCGCGCGCTTGA..... 519
327 ArgThrProProArgAlaSerProThrGlyThrProSerArgAlaSerPr 343
519 ..... 519
343 oThrGlyThrProSerArgAlaSerLeuThrGlySerProSerArgAlas 360
518 .....TCCGCGCGCTTCTTCGATTAATGACCGTAGGCTGC 483
360 eLeuThrGlyThrProSerArgAlaSerLeuLeuGlyThrProSerArg 376
482 GCAGCCAGCGGATTTGTCCTCATTCGATTCAGAAATGCGCAACGCTC 433
483 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
377 AlaSer.....LeuLeuGlyThrProSerArgAlaSerLeuThrGlyTh 391
432 GGCATTCAGCGCAGCAATTTTCGTAACGAGCGGCTGCGACGACATCG 383
433 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
391 rProProArgAlaSerLeuThrGlyThrSerSerThrAlaSerLeuThra 408
382 ACAACCGGATTTGATGATGTTGCGCGCGCACCTTCCTGCGCTTAAGTTT 333
383 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
408 rGThrProSerArgAlaSerLeuThrArgThrGlnSer..... 420
332 GCCAGCGCTTACGTGCTAGCGCTTCGAACCTCGATTCGTCGCTTC 283
421 .....SerSerLeuThrArgThrProse 429
282 AACGGCAATCAGACTGACTGATGATGCGGCTTTCGCCAGCGTGAATCG 233
429 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
429 rMetAlaSerLeu.....ThrArgThrProProArg.....A 440
232 CGGCGATTTTG.....CCTGAAGCGCGCGCATTAACACCAACGCGCC 192
233 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
440 lSerLeuThrArgThrProProArgAlaSerLeuThrArgThrPro 456
191 GATTCCTTTTGTCAACACGACCTTGCGCTTTCGACGCGCATCGCC 142
457 ArgAlaSerLeuThrArgThrProProArgAlaSerLeuThrArgThrPr 473
141 TTCCTGACTTCATGAGGCGCATACCGGCAATATTCCTCG..... 99
473 oSerMetValSerLeuLys...ArgSerProSerArgAlaSerLeuThra 489
98 .....CCAAAGCAACGCGACT.....TCGCTA 78
489 rGThrProSerArgAlaSerLeuThrMetThrProSerArgAlaSerLeu 505
77 ATGGCGGCGCGCTGTAACGCGCTTCGCTCGGCTGCGC 39
506 ThrArgThrProSerThrAlaSerLeuThrGlyThrPro 518
seq_name: p1r2:147141

```

seq documentation block:
 gastric mucin (clone PGM-2A) - pig (fragment)
 C:Species: Sus scrofa domestica (domestic pig)
 C:date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 03-Nov-2000
 C:Accession: 147141; 555315
 R:Turner, B.S.; Bhaskar, K.R.; Hadzopoulos-Ciadaras, M.; Specian, R.D.; Lamont, J.T.
 Gastroenterology 106, 200, 1994
 A:title: Pig gastric mucin: isolation and characterization of a cDNA clone with a nov

A:Reference number: I47141; MUID:94102478
 A:Accession: I47141
 A>Status: Preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-528 <TR>
 A:Cross-references: EMBL:U10281; NID:9915205; PIDN:AAC48526.1; PID:9915208
 R:Turner, B.S.; Bhaskar, K.R.; Hadzopoulou-Cladaras, M.; Specian, R.D.; Lamont, J.T.
 Biochem. J. 308, 89-96, 1995
 A:Title: Isolation and characterization of cDNA clones encoding pig gastric mucin.
 A:Reference number: S55315; MUID:95275264
 A:Accession: S55315
 A>Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 1-528 <TR>
 A:Cross-references: GB:U10281; NID:9915205; PIDN:AAC48526.1; PID:9915208
 C:Superfamily: pig submaxillary mucin

Alignment scores:
 Quality: 113.50 Length: 458
 Ratio: 0.493 Gaps: 19
 Percent Similarity: 50.218 Percent Identity: 20.961

Alignment_block:
 US-09-303-518D-125/rev x I47141 ..

Align seg 1/1 to: I47141 from: 1 to: 528

```

1343 CAGCCTTCCTTCGATG.....GTTTCAGACATTGCGCAACAG 1303
149 GlnProSerSerSerSerSerProIleSerSerThrVal..... 162
1302 CGGCGCGATTCGATTTGCGCGGCGAGACAGATCGACAAAGCAGGT 1253
163 .....SerValGlnProSerSerSer.....SerAlaProT 174
1252 CTTCTGCTGCAATTCAGACCAACCAATGCTGCGCGCTG...TCGTA 1206
174 hrThrSerAlaThrSerValGlnProSerSerSerSerProIle 190
1205 TCGCGCAGATTAATCGCGCAAGAGCGGCGAGATATCAAGG 1156
191 SerSerThrValSerValGlnThrSerSerSerSerSerValProThr 207
1155 CAGCAGCGCTGTAAGTACCAATCGGACCATGCGCGGCTGCGCGCT 1106
207 rSerThrThrSerValGlnPro.....SerSerSerSerSerValProT 222
1105 TGACGCTGTGTGAACCTGAAGATTGTTTTCAGAAATGCGCGAG 1056
222 hrThrSerAlaThrSerValArgSerSerSerSerSerThrPro... 237
1055 GTTGTACGCGCTGATGAGTATTGTGCGGCTGCGGCGCAACCGCGAA 1006
238 .....IleProSerThrThrSerValGlnProSe 247
1005 CAGCTCTTGTGCGGCT...TCTTCAGTAACGAATCTGATGTGTG 959
247 rSerSerSerSerAlaProThrThrSerAlaThrSerVal..... 260
958 AGGTCCTCAATATATGTCGCGCTGTGTATCGCGCGTTCATATAC 909
261 ..GlnProSerSerSerSerThrProIleProSerThrThrSerVal 276
908 GAACGGAATACAGCGGTGTGTGTAACCAATTCGCCCGCGATAT 859
277 GlnPro.....SerSerSerSerSerAlaProThrThrSe 288
858 TTGCGATCTTTCGACCAACCAACGATACGACAGCGCGGTGTGTA 809
288 rAlaThrSerValGlnProSerSer.....S 297
808 CTTGAGAACCACTAGGCAATACAGCGCGTGTGTTGACAGCGCGCT 759

```

```

297 ..... 311
297 rSerSerProIleSerSerThrIleSerVal.....GlnProSer 311
758 GCAACAACAGCGCCAAATGTAATTCATCTGTAATGATGTCACAC 709
312 SerSerSerSerProThrThrSerThrThrSerValGlnProSer 328
708 GGTTCGTCGCGCGCGCGCTGCAATGAATGATGCGTCCACTCA 659
328 rSerGlySerAlaProThrThrSer.....AlaThrSerValG 341
658 ACCGCGAGATCGGCGCGCAATTCATGTTTCGATGTTGCGACA 609
341 InProSerSerSerSerSerProIleSerSerThrIleSerValG 357
608 TTTTCAGACGCGACGTCGCGCGCGCTTACAACATGATGTTGCG 559
358 ProSerSerSerSerSerProThrThrSerThrThrSerVal.... 372
558 TTGCGTCAACGCGCTCATACACAGCGCGCTTGAATCCTCGCG 510
373 .....GlnProSerSerSerGlySerAlaP 381
509 .....GTTCTTTGATATATACCGTAGGTCGCGACCGCGATG 465
381 rThrThrSerAlaThrSerValGlnProSerSerSerSerValPro 397
464 TCCATGCGATTACGACAGATGCGCAACGCGCGCATCGACGAGAT 415
398 ThrThrSerAlaThrSerValArgSerSerSerSerSerThrProI 414
414 TTTGCGAAGGAGCGGTCGCGACGCGATCCACAACCGGATTGATCA 365
414 eProThrThrThrSerValGlnProSer..... 423
364 GGTTCGCGCGCACTTCTTCGCGCTTAAATTGTCAGCGCTTCAGT 315
424 .....SerSerSerSerValProThrThrSerAlaThrSerVal 436
314 TAGCCTCAACTGATTTGCTG...TTGCTTCAACGCAATCAGAC 268
437 GlnThrSerSerSerSerSerThrProIleProSerThrThrSerValG 453
267 TGACTGAAGTACGCGCTTTCGCGCACGCGTGAATGCGCGATTTG 218
453 nProSerSerSerSerSerAlaPro..... 461
217 AAGCGGCGCAGTAACACACGCGCGGATTTCTTTGTCTTCAACAG 168
462 .....ThrThr.....SerAlaThrSer 467
167 ACTTGGCCTTTTTCAGCGCAATGCGCTTCTGACTTCATCGAGGCG 118
468 ValGlnProSerSerSerSerSerProIleSerSerThrIleSerVal 484
117 CATACGCGATATTCCTGCGCAACGCGACTTCGATATG..... 75
484 GlnProSerSerSerSerSerProThrThrSerThrThrSerValG 501
74 .....GCCGCGCGCTGCTAAGCGTTGCTGCTCGGCTG 42
501 InProSerSerSerGlySerAlaProThrThrSerAlaThrSerValG 517
41 CCGCGATGCGGCAAGTTAGACCT 18
518 ProSerSerSerSerSerProPro 525

```

seq_name: p1r2:F75196

seq_documentation_block:
 pyruvate, water dikinase (EC 2.7.9.2) PAB0057 - Pyrococcus abyssi (strain Orsay)
 N:Alternate names: phosphoenolpyruvate synthase
 C:Species: Pyrococcus abyssi

C>Date: 11-Feb-1993 #sequence_revision 11-Feb-1993 #text_change 18-Nov-1994
 C:Accession: C39135
 R:Holmes, M.L.; Dyall-Smith, M.L.
 J. Bacteriol. 173, 642-648, 1991
 A:Title: Mutations in DNA gyrase result in novobiocin resistance in halophilic archaea
 A:Reference numbers: A39135; MUID:91100352
 A:Accession: C39135
 A>Status: preliminary; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-437 <HOL>
 A:Cross-references: GB:M38373

alignment_scores:
 Quality: 112.00 Length: 450
 Ratio: 0.626 Gaps: 23
 Percent Similarity: 39.778 Percent Identity: 24.222

alignment_block:
 US-09-303-518D-125 x C39135 ..

Align seg 1/1 to: C39135 from: 1 to: 437

```

36 CGCGGGACAGCGGACGCGT.....TTACGACGGCC 70
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
94 ArgAspGlnAspGlnAlaProAlaArgHisLeuAlaArgAlaGly 110
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
71 CGCGCATACCGAAGTGGCGTGGCGAAGATATGCCGATATCGC 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
110 GtGtGtAlaHisArgGlyArg..... 116
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
121 CCCCGATGAAGTCAAGAGAGGATCGGTCAAAAAGGCCAGTGT 170
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
117 .....AsnAlaGlnHisAlaGlyAspArgAlaProGlyValAspSerArg 131
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
171 GTTGAAGACAAAAGATCCGGCGTGTGTTACTGCGCGCCG 215
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
132 LeuArgGlnGlnHisGlnHisProArgGlyArgHisAlaSerAspArg 148
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
216 .....TTCAAGCAAAATCGCGGATTCACCGTGGC 246
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
148 GlnAspGlyAlaHisProArgAlaArgLeuArgGlnProArgH 165
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
247 GAAAGCGCGTACTTCAGTCAGTCGATGTCGGTGAAGGCA 290
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
165 AlaAlaGlyArgPro.....ArgArgArgGlnProPro 175
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
291 CGACGAATGAGATTGAACGCTACGACCTGAGCGCTGCAAACTAA 340
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
176 ArgArgGlyArg.....SerArgGlyThrHisArg 185
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
341 GCGGCGA...AGAAGTGGCGCGCACTGATCCATCCGTTTGGAGCT 387
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
185 GtGtGtHisLeuAlaGlnAlaProArgPro...AlaValArgGlyProAsp 200
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
388 GCGCTCGGACCGCGTTCACGAAATTCCTGCGCGATGCGAGCG 437
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
201 .....GlnAspGlnAlaArgGlnLeuArgGly 209
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
438 GTTCGGCAT.....CTTCG 451
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
210 ProArgHisArgArgGlnAlaArgHisProProThrAlaArgAspValLeuArg 226
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
452 TCAATGCGATGGACACCAATCGCGTGGCGCCACCTACGCTATTATC 501
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
226 GtGtGtGlnProGlnHisGlyAsp.....GlnHisHisL 237
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
502 AAAAGACCGCGGACGATTCAAACGCGCGCTGTGATGAGCGCGT 551
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
237 euglGtGtArgArgGly...ArgProArgProGlnGtArgGlnAlaGly 252
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
552 GACCGACGCAAAATCCATGTTTGAAGGACAGTGGCGGACGCGTCCGT 601
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

253 ArgGlyAlaHisProProGlnVal.....ArgAlaArgIleGly 265
602 CTGAAATGCTGCCACATCATGAAACATGATTCGGCGCGCCGATCC 651
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
265 rLeuAlaAlaGlyGlnAlaArgGlyLeuProGlnProArgProLeuGly 282
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
652 GCGGCTTGGATGGACGCGCATTCATTTCATGACCGCGTGGCGGAA 701
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
282 AlaArg.....ThrValHisArgGlyGlnArgLeuArg 292
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
702 TAAACCGTGGACCATCATATGACAGTGAATTCATTCATTCGCGCT 751
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
293 ...GlyArgValGlnAlaGlyProArgProGlnValProGlyAspPh 308
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
752 TGTTCGCAACAGCGCGTGTGAACACGCGCGGTGATTCCTAGTGTGT 801
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
308 eAlaProGlnGlyGlnAspSerGlnArgGlnArgProProArgProH 325
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
802 TCTCAAGTCAACAAACCGCGCTTGGCTACCGTTTGGTGGCAAGT 851
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
325 AsSer..... 326
852 ATCCGAATTAATTCGCGGCAATGTTGACACAGCAACCGCTGATTT 901
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
327 .....ArgLysArgArgAsp.. 331
902 CGGCTTCGATTAAGCGCGCGATTCACAAAGCGCGACGATTTATGG 951
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
332 ...ThrGlyAlaHisArgHisArgHisArgArg..... 342
952 GGACGCTACCAATCAGATTTCCTTATGCAAGAAAGCGCGAGCAAGA 1001
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
343 .....ArgArgArgValArgHisArg 349
1002 GCTGTTCGCGTGGTGGCGCGACCGGACAAATTCATCCACGCGTA 1051
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
350 GlnGlyAlaLeu..... 353
1052 CAACCTCGCGCATTTCTGAAACAAACACTTCAGTTCAACACAGCC 1101
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
354 ...ProAlaAlaHisProAspAspArg.....A 362
1102 GTCAAGCGCGCGGACCGCGCGATGTCGATGTTACTTACGAGCGGT 1151
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
362 rGArgArgArgArgArgAlaHis.....Pro 370
1152 GATGCCCTTGATATCCCTGCCACCGCTCTTTGCGGATTAATCGTGC 1201
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
371 AspAlaAlaAlaTyrAlaSerValProAlaHisAlaProAlaHisArg 387
1202 GCGATACCGACAGCGCGGACGATGTTGCTTGAATGGAGAGAGA 1251
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
387 rArgLeuAlaValAlaArgLysSerThrAla.....AlaValProArgProL 402
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1252 GACCTCGCTTGTGACGCTGCTGCGCGGCAATGCAATAGCGGCC 1301
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
402 eupProArg...GlnHisLeuArgArgAspGlyArgGlnAlaGlyPro 417
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
seq_name: p1r2:152257
seq_documentation_block:
epistatin - mouse
C:Species: Mus musculus (house mouse)
C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 05-Nov-1999
C:Accession: I52257; 152210
R:Yos, H.L.; De Vries, Y.; Hilkens, J.
Biochem. Biophys. Res. Commun. 181, 121-130, 1991
A:Title: The mouse epistatin (Muc1) gene and its promoter. Rapid evolution of the rep
A:Reference number: I52257; MUID:92068178
A:Accession: I52257
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-631 <RES>

```


A:Cross-references: GB:AE001863; GB:AE001825; NID:96460670; PIDN:AAF12376.1; PID:9646067
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DRA0268
 A:Map position: 2

alignment_scores:
 Quality: 111.50 Length: 301
 Ratio: 0.945 Gaps: 15
 Percent Similarity: 39.203 Percent Identity: 26.246

alignment_block:
 US-09-303-518D-125 x C75580

Align seg 1/1 to: C75580 from: 1 to: 376

```

612 TGGCAACATCGAAMACACATGATTCGGCGCCGATCCGCGTTGA 661
|||||
32 CysATGHisATGProAla...LeuATG1yAspAspSerAlaATGLeuAl 47
662 GTGGCGCGACATTCATTCAT..... 683
47 ag1yHisAlaHisHisThrHisPheGlyAlaAlaGlyATGThrGlyThrG 64
684 .....CGAGCGCGTGGCGCGCGGATTA 704
64 LyATGProAlaAlaGlyProAlaGlyGlnSerGlyAlaGlyLysPro 80
705 AACCGTGTGGACATTCATTCATTCATTCATTCATTCATTCATTCAT 754
81 ATGATGValAlaHis..... 85
755 TTGCAACGCGCGTGTGAACACGCGCGGATTCGCTAGCGTTCT 804
86 .....ProAlaLeuLys 90
805 CAAGTCAACAACCGCGCTTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 854
90 erg1yGlyATG...AlaAspATGAlaGlyATGAlaGlySerAlaAlaGly 105
855 GCAAAATTCGTCGCGCGATTCGTCGTCGTCGTCGTCGTCGTCGTCG 904
106 AlATrGTyATrGTu1LeuGlyProAlaATG...ProATrGlyAspAT 121
905 GTTCGATATTCGACGCGCGATTCACAAAGCGCGACGATTCGTCGTCG 954
121 gProATrValSerATrATG...LeuATrAlaAlaAlaATrLeuAlAT 136
955 CGCTA..... 959
136 hTrLeuATrGlyAlaGlyAsnLeuATrGlyATrATrGlySerAlaAlaATr 152
960 .....CSACAATCAGATTCGCTTCGTCGTCGTCGTCGTCGTCGTCG 997
153 ATrATrGAspAlaATrProAla...ATrATrATrATrATrATrATrATr 166
998 AAGAGCTGTTCGCTGCGT..... 1016
167 .....AlaATrGLeuGlyATrGlyHisLeuATrSerAlaATrAlaLeuAl 181
1017 .....TGCGCGCGACGC 1028
181 hATrHisAlaATrAspValATrATrATrATrATrATrATrATrATrATr 197
1029 GGAACAATATCTCCATCAGCGCGTCGTCGTCGTCGTCGTCGTCGTCG 1078
198 g1yGlnATrATrATrATrATrATrATrATrATrATrATrATrATrATr 214
1079 AACCTTCAGATTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1128
214 gValATrAlaGlyLeuLeuAspValATrATrATrATrATrATrATrATr 231

```

```

1129 CCGATGTGACTTACGACCGCGTGCATCCCTTGATATCTGCCAC... 1175
231 LysHisGlnHis...SerAlaATrATrATrATrATrATrATrATrATr 246
1176 .....CCTGCT.....TTGGCGGATTCATTCGTCG 1201
247 GlnATrATrGlyProAlaATrGlyGlySerGlyAlaATrThrAlaATrAT 263
1202 GCGATACCGACACCGCGCGACGTCGTCGTCGTCGTCGTCGTCGTCG 1251
263 gATrHisATrATrATrATrATrATrATrATrATrATrATrATrATrAT 273
1252 GACCTCGCTTGTGACAGCTTCGTCGTCGTCGTCGTCGTCGTCGTCG 1301
273 LyProATrGlyAlaValAlaProAlaLeuATrGlySer...AspAspATr 288
1302 GCT 1304
289 Ala 289
seq_name: p1r2:T45134

```

seq_documentation_block:
 hypothetical protein [imported] - Microbacterium ammoniaphilum (fragment)
 C:Species: Microbacterium ammoniaphilum
 C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
 C:Accession: T45134
 R:Strlebel, H.M.; Seeber, S.; Jarsch, M.; Kessler, C.
 Gene 172, 41-46, 1996
 A:Title: Cloning and characterization of the Mami restriction-modification system fro
 A:Reference number: 229233; MUID:96257250
 A:Accession: T45134
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-529 <STR>
 A:Cross-references: EMBL:X79027; NID:9984667; PIDN:CAA55649.1; PID:g1679831
 A:Experimental source: ATCC 15354

alignment_scores:
 Quality: 111.00 Length: 526
 Ratio: 0.491 Gaps: 29
 Percent Similarity: 42.966 Percent Identity: 24.715

alignment_block:
 US-09-303-518D-125 x T45134

Align seg 1/1 to: T45134 from: 1 to: 529

```

21 TCTAACTGCCCATCGCGGCGACGACGACGACGACGCGTTTACGA...CG 67
49 AlATrGProGlyATrGlyAlaGlyProATrGlyThrc1yThr1LeG1uProAT 65
68 GCCCGCATTCACGAGATTCGCTGTCGTCGTCGTCGTCGTCGTCGTCG 117
65 gG1ySerHisATrATrATrATrATrATrATrATrATrATrATrATrATr 77
118 CGCCCTCGATGAATGCAAGGAGCGGATCGCTCAAAAAGGCGCAAT 167
78 .....ATrGlnValHisProLeu 83
168 GCTGTTTGAAGACAAAGAATCGCGCGTGTATTCGCGCGCGCTT 217
84 GlnATrProHisGlnGluProGlyAlaATrGlyGlu.....LeuProVa 98
218 CAGGCAAAATCGCGCGATTCACCGTCGTCGTCGTCGTCGTCGTCGTCG 267
98 lATrAspATrATrATrATrATrATrATrATrATrATrATrATrATrAT 115
268 GTGCTGATTCGCT..... 281
115 lATrAspAlaATrATrATrATrATrATrATrATrATrATrATrATrATr 131

```

```

282 .....TGAGGCAAGCAAGATTCGATTTGAACSTGACACSTG 322
      |||.....|||
132 G1yAlLeuArGlnIsArGArGlnIsArG..... 141
      |||.....|||
323 AAGGCGTGGCAAACTTAAGCG.....CGAAGAAGTGGCGGCGCAACSTG 366
      |||.....|||
142 ALArGArG1yAlLeuArGlnIsArGArGlnIsArGArGlnIsArG 158
      |||.....|||
367 ATGCAATCGGTTTGTGACTGCGTGGCGGCAACSTGCGTTCAGCAAAAT 416
      |||.....|||
158 IsPrArGArG1yAlLeuArGlnIsArGArGlnIsArGArGlnIsArG 173
      |||.....|||
417 TCGTGGCGTGGCGGCAACSTG.....GTTCG 442
      |||.....|||
174 .....ArGArGArGArGArGArGArGlnIsArGArGlnIsArG 188
      |||.....|||
443 CGATTCGTGCAATCGGATTCGACACCAATCGGCGTGGCGGCAACSTG 488
      |||.....|||
188 gHIs.....G1yArSArSglnHIsArGAlAsPrArGArGArGPrArGAr 203
      |||.....|||
489 .....TACGTCATTA.....CAAGA 506
      |||.....|||
203 sPrArGArGAlAg1yHIsArGlnIsArGlnIsArGlnIsArGlnIsAr 219
      |||.....|||
507 AGCCGC.....CGAGATTCGAACGCGCGCGTGGTGGATTCGA 544
      |||.....|||
220 ALArArGPrArGlnIsArGArGlnIsArGlnIsArGlnIsArGlnIsAr 234
      |||.....|||
545 GCGGTTTGGCAAGCAAGATTCGATTTGAAGCGACSTGGCGC..... 590
      |||.....|||
235 .PrArAlAg1yArG.....HIsPrArAlLeuG1yArSglnIsArSglnIsAr 249
      |||.....|||
591 .....AGACGTGGCGTGGCAAAATGC 611
      |||.....|||
249 rGAlArGlnIsArGAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAl 265
      |||.....|||
612 TGCGCAATCGCAAACTTAAGCGCGCGCGCGCGCGCGCGCGCGCGTTCGA 661
      |||.....|||
266 LeuGlnArGArGArG.....G1yArGAlAsPrG1ySArGArGAlAg1 279
      |||.....|||
662 GTGGCGGCAATTCGATTCGACACCGGT.....CGCG 696
      |||.....|||
279 nG1yArGAlArGArGAlArGArGAlAg1yGlnIsArGlnIsArGlnIsAr 296
      |||.....|||
697 GCGAATTAAGCAAGCGTGGCGACATTCGATTCGAAGTGAATTAACATTCG 746
      |||.....|||
296 rGlnIsArGlnIsArGlnIsArGlnIsArGlnIsArGlnIsArGlnIsAr 304
      |||.....|||
747 CGG.....TT 751
      |||.....|||
305 PrArArGPrArGArGlnIsArGlnIsArGlnIsArGlnIsArGlnIsAr 321
      |||.....|||
752 TGTTCGAACAGCGCGTTCGAACAGCGCGCGCGCGCGCGCGCGCGCGTTCG 797
      |||.....|||
321 yAlAlG1yAlArGPrArAlArGPrArGArGlnIsArGlnIsArGlnIsAr 338
      |||.....|||
798 .....TGGTTCGCAAGTGAACAAACAGCGCGCGTTCGCGTTCGCGTTCG 836
      |||.....|||
338 lAsArPrArGlnIsArGlnIsArGlnIsArGlnIsArGlnIsArGlnIsAr 354
      |||.....|||
837 TTTGGGTGGCAAGATTCGCAAACTTAAGCGCGCGCGCGCGCGCGCGCGCG 886
      |||.....|||
355 GlnG1yLeuGlnIsArGlnIsArGlnIsArGlnIsArGlnIsArGlnIsAr 365
      |||.....|||
887 ACAACCGCGCGATTCGCGCGTTCGCGTTCGCGTTCGCGTTCGCGTTCG 930
      |||.....|||
365 gHIsPrArGlnIsArGlnIsArGlnIsArGlnIsArGlnIsArGlnIsAr 382
      |||.....|||
931 CAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 980
      |||.....|||
382 eulArGArGlnIsArGlnIsArGlnIsArGlnIsArGlnIsArGlnIsAr 991
      |||.....|||
981 CGAAGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1030

```

```

      |||.....|||
392 ArGArGlnIsArGlnIsArGlnIsArGlnIsArGlnIsArGlnIsAr 404
      |||.....|||
1031 ACMAATATCTGATCGACGCGTTCGACACCGCGCGCGCGCGCGCGCGCGCG 1080
      |||.....|||
404 yGlnG1yLeuArGlnIsArGlnIsArGlnIsArGlnIsArGlnIsArGlnIsAr 420
      |||.....|||
1081 CTCTTCAATCTGATCGACGCGTTCGACACCGCGCGCGCGCGCGCGCGCGCG 1130
      |||.....|||
421 LeuAlAlLeuArGlnIsArGlnIsArGlnIsArGlnIsArGlnIsArGlnIsAr 437
      |||.....|||
1131 GATTGGTACTTTCGACGCGCGTTCGATTCGATTCGATTCGATTCGATTCG 1180
      |||.....|||
437 utPrMetThrSerG1yArGlnIsArGlnIsArGlnIsArGlnIsArGlnIsAr 452
      |||.....|||
1181 TTTGGCGCGCGTTCGATTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1212
      |||.....|||
452 eucysArGArSArSerValGlnSerHIsLeuArSArValHIsThrThrArGAla 468
      |||.....|||
1213 .....AGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1250
      |||.....|||
469 AsPrAlAserThrSer.ArGArGlnIsArGlnIsArGlnIsArGlnIsAr 484
      |||.....|||
1251 AGACSTTCGTTTTCGACGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1300
      |||.....|||
484 eulArGlnIsArGlnIsArGlnIsArGlnIsArGlnIsArGlnIsArGlnIsAr 500
      |||.....|||
1301 CGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1324
      |||.....|||
501 SerThrSerArGlnIsArGlnIsArGlnIsArGlnIsArGlnIsArGlnIsAr 508
      |||.....|||
seq_name: p1r2:T43311
seq_documentation_block:
Fatty acyl-CoA synthase (FC 2.3.1.86) beta chain - fission yeast (Schizosaccharomyces
N:Alternate names: Fatty acid synthase beta chain
C:Species: Schizosaccharomyces pombe
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
C:Accession: T43311
R:Niwa, H.; Katayama, E.; Yanagida, M.; Morikawa, K.
Protein Expr. Purif. 13, 403-413, 1998
A:Title: Cloning of the fatty acid synthetase beta subunit from fission yeast, coexpr
A:Reference number: Z22415; M01D:98359895
A:Accession: T43311
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2073 <NTM>
A:Cross-references: EMBL:AB010274; NID:94115479; PIDN:BAA36384.1; PID:94115480
C:Genetics:
A:Gene: fas1
C:Superfamily: Yeast fatty-acyl-CoA synthase beta chain
C:Keywords: acetyl-CoA; acyltransferase; carbon-oxygen lyase; coenzyme A; fatty acid
alignment_scores:
Quality: 110.50 Length: 465
Ratio: 0.521 Gaps: 25
Percent Similarity: 45.591 Percent Identity: 20.215
alignment_block:
US-09-303-518D-125 x T43311 ..
Align seg 1/1 to: T43311 from: 1 to: 2073
217 TCAGGCAAAATCGCGCGATTCACCGTGGCGAAGCGCGCGCGCGCGCGCGCGCG 266
      |||.....|||
948 ThrG1yLysLeu.....MetArGArGlnIsArGlnIsArGlnIsArGlnIsAr 962
      |||.....|||
267 AGTC.....GTGATTCGCGTTCGAAGCAAGCAAGCAAGTTCGAG..... 303
      |||.....|||
962 pValG1yLysThrThrLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 979
      |||.....|||
304 .....TTGAACGCTACGACCTGAGCGCGTGC 333

```



```

334 AACTTAAGCGCGGAGAGAGTGGCGGCAACCTGATCCAAATCCGGTTTG 383
      ::::: :::::
996 Aspleusntrhrlaspralagl.....Pheh 1005
384 GACTGCGCGCGGCGCGCGTTCGTTACAGAA.....ATTCTG 421
      ::::: :::::
1005 etylserleucylserasnprophegluysprovalprophellepro 1022
422 CCGTCGAT..... 429
1022 laileasprathrpheluphethrphelysaspserleutrpel 1038
430 GCCGAGCGCGTTCGCCATCTTCGTAAT.....GGAT 461
      ::::: :::::
1039 Sergluaspralalalavalalvalalgluaspralglarvalalal 1055
462 GACACCAATCCGCGTGGCGGACCCAGCGTCATATCAAGAGCG 511
      ::::: :::::
1055 eleuglndlyprometalalalalysnhsertthlysalnslupro 1072
512 CCGAGATTTCAAAAGCGCGCTGTTGATTGAGCGGTTTGACCGAGCG 561
      ::::: :::::
1072 lalysglu.....leuaspgllyleasnsluphr 1082
562 AAAATC...CATGTTTGAAGCA.....GT 596
      ::::: :::::
1083 HisileglnhlsPheilelyspheTyrAlaglyaspgluLysI 1099
597 GCGGCTGAAATGCTGCCAATCGAACATGATTCGCGCGCGCG 646
      ::::: :::::
1099 eprolle.....Valglutyrphelgllyvalp 1109
647 ATCTGCGCGTTGAGT..... 663
      ::::: :::::
1109 roprovalasnvalserhlsysserleugluservalserValthrslu 1125
664 .....GGACGCGCATTTATTCATTCGAGCGCGCGCGCAATAAAC 707
      ::::: :::::
1126 Glualaglyserlyvaltyrlylsleuprogliulleglseraseral 1142
708 CCG..... 711
      ::::: :::::
1142 aleuproserylsylsleutrphegluleuAlaglyproglutyrT 1159
712 .....TGCAC 717
      ::::: :::::
1159 htrphearalalalepherhtrhglarvalalalysglutrypls 1175
718 ATCAATTTACAGATGTAAATTACATTTGCGGTTTGAACAGCGCG 767
      ::::: :::::
1176 leuglhnhsasnproval.....Argargllephealapro..... 1187
768 TCTGAACACCGAGCGCGGATTCGCTAGGTGCTTCAAGTCAACAAC 817
      ::::: :::::
1188 .Argtyrgllyglalnargalavalalysgllyasasasnspthval 1204
818 CCGCGCTCTTGGCTACC.....GTTTGGTGGGAAGTA 852
      ::::: :::::
1204 algluleutyrglutrhglnserylAsnTyrValleuAlalaaglu 1220
853 TTGCAATTTACTGCGGCGAATGGTTGGTACACAGACCGCGTGAATTC 902
      ::::: :::::
1221 Ser.....Tyraspgllyglutrhlevalalase 1230
903 CCGTTCCGATTTGAACCGCGCGATTTACACAGCGCGCAC...GATTAT 949
      ::::: :::::
1230 rmetphelgluasnrgrasnalaleuyslysglvalhlsleuasphe 1247
950 TGGAGCGCTAC.....CACATGAGATTTCCGTTATTCGA 984
      ::::: :::::
1247 euphelystyrgluProserAlaglyTyrSerProvalSergluileu 1263

```

```

985 GAAGCGCGGACGAAAGAGCTG.....TTCGCTGGTT..... 1017
      ::::: :::::
1264 Aspglyluprasnaspargllylelyshspherlytrrpalaleutrphegl 1280
1018 ....GCGCGGACCGGACAAATACTCATCGACGCGTACACCTCGGCC 1063
      ::::: :::::
1280 yglugluprotyrprogluasnalaserllethrasprthrphetrhglp 1297
1064 ATTCTCTGAA.....AACAAACCTTCAGTTCAACACAGCGCGC 1104
      ::::: :::::
1297 rogluvalthrvalthrgllyasnmetllegluasphelcysarghtrval 1313
1105 AACGCGCGGACCGCGCATGTGTCGATTCGATTCAGACCGCGTGA 1154
      ::::: :::::
1314 GlyAsnhsasnslualatyrthrlysalargalalalearglyargmetAl 1330
1155 GCGCTTGAT.....A 1165
      ::::: :::::
1330 aprometasphealalilevalalglutrpglalalalethrlysAlat 1347
1166 TCCTGCCACCCGCTTTGCGGATTA.....ATCGTCGCGATACC 1209
      ::::: :::::
1347 lepheprolysalalileaspgllyaspheleuargleuvalhlsleuser 1363
1210 GACAGCGCGGCGCATTTGGCTTGGATTGACGAGAAAGAC 1254
      ::::: :::::
1364 Asnserphearglmetvalgllyserhlsleuemetgllyasp 1378

```

seq_name: p1r2.T28969

```

seq_documentation_block:
  hypothetical protein T23H2.1 - Caenorhabditis elegans
  C:Species: Caenorhabditis elegans
  C:Date: 29-Oct-1999 #sequence revision 29-Oct-1999 #text_change 17-Mar-2000
  C:Accession: T28969
  R:Wamsley, P.; Bradshaw, H.,
  submitted to the EMBL Data Library, January 1997
  A:Description: The sequence of C. elegans cosmid T23H2.
  A:Reference number: Z20549
  A:Accession: T28969
  A:Status: preliminary; translated from GB/EMBL/DBJ
  A:Molecule type: DNA
  A:Residues: 1-1847 <MAN>
  A:Cross-references: EMBL:U80033; PDB:AAQ48199.1; GSPDB:GN00019; CESP:T23H2.1
  A:Experimental source: strain Bristol NZ; clone T23H2
  C:Genetics:
  A:Map position: 1
  A:Introns: 52/2; 102/1; 223/3; 311/1; 370/3; 798/3; 1107/1; 1230/2; 1282/3; 1342/2; 1
  C:Superfamily: rat nuclear pore protein gp210

```

```

alignment_scores:
  Quality: 110.00      Length: 511
  Ratio: 0.464        Gaps: 26
  Percent Similarity: 46.380      Percent Identity: 22.114

```

alignment_block:

```

US-09-303-518D-125 x T28969 ..
Align seg 1/1 to: T28969 from: 1 to: 1847

4 ATTAATCAAAAAAGGTCTAAACCTGCCATCGCGGCGACAGCGAGCA 53
      ::::: :::::
818 lletthvalysgllyleuasnval.....gluprolalAr 830
54 ACCGCTTTCAGACGCGCGCGCATTTACCGAAGTCCGCTGGT..... 96
      ::::: :::::
830 gleuvalleutrpasnglualalavalserlyseljythrphetrhlerng 847
97 .....GCGGAGAAATATTCGCGGTATGCGCGCTCGATGAAGTCAAG 138
      ::::: :::::
847 lygllysercllythshhsvalaspasnleuprotthrsaspserPro 863

```



```

428 Ser.....ThrLeuGlySerSerSerThrIleLeu 438
465 GTCCATCGCATGTCAGACAGATGCGCAACCGCTCGCATCGACGACGAGA 416
      ||||| ..... |||||
438 lSerThrThrIleThrPro...SerThrProLysValSerThrIleThrL 454
      ||||| ..... |||||
415 TTTTGCTGAACGAGACGGGTGCGCAGTCACAAACCGATTTGGATC 366
      ||||| ..... |||||
454 eu.....SerGlnSerProThrPro..... 460
365 AGATTGCGGCGCATCTTCGCGCTTAACTTGGCAGCGCTTCAGTGC 316
      ||||| ..... |||||
461 .....ThrSerThrProLeuValValSerSerSer..... 471
315 GTACGCTTCAACCTGATTTGCTGCTGCTTCAACGCGCAATCAGCAGT 266
      ||||| ..... |||||
472 .....SerGlySerSerSerThrValValThrSerThrIleThrPro. 485
265 ACTGAAGTACGCGCTTTTCGCGCAGGTGAATCGCGCATTTTGGCTGAA 216
      ||||| ..... |||||
486 .....SerThrGlnGly.....ValProThr 492
215 GCGGCGCGCATTAACACACGCGCGATTTTGTGCTTCAACAGCAGAC 166
      ||||| ..... |||||
493 SerThrSerAsnGlnProThrProSer.....ThrSerAsnProThr 506
165 TTGGGCT.....TTTGTAGCGCATGCGCTTCTTGAATTCATCGAGC 122
      ||||| ..... |||||
506 rThrProLysSerThrValThrAlaSerProSerThrThrGlyAlaThrS 523
121 GCGCGCA.....CCGCAATATCTTCGCCAAGCAGCGACTGGGTATG 75
      ||||| ..... |||||
523 erThrAlaSerProSerThrIleThrSerSerAlaProThrSerGlnSer 539
74 GCCGCGCGCTGTAAACGCGCTTCTCGGTTCGCCGATGGCGCAGATT 25
      ||||| ..... |||||
540 HisSerProSerSerThrMetThrSerThrValProValThrSerThrph 556
24 T 24
556 e 556

```

seq_name: pir2.T34434

seq_documentation_block:
 hypothetical protein K06A9.1a - Caenorhabditis elegans
 C.Species: Caenorhabditis elegans
 C.Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 18-Feb-2000
 C.Accession: T34434
 R.Gelisel, C.: Gatlung, S.
 Submitted to the EMBL data library, December 1996
 A.Description: The sequence of C. elegans cosmid K06A9.
 A.Reference number: 221525
 A.Accession: T34434
 A.Status: preliminary; translated from GB/EMBL/DBJ
 A.Molecule type: DNA
 A.Residues: 1-2232 <GBI>
 A.Cross-references: EMBL:U08046; PIDN:AC70890.1; GSPDB:GN00028; CESP:K06A9.1a
 C.Experimental source: strain Bristol N2; clone K06A9
 C.Genetics: CESP:K06A9.1a
 A:Map position: X
 A:introns: 38/1; 75/3; 103/3; 132/2; 158/2; 222/1; 1088/1; 1367/1; 2039/1; 2049/1; 2075/1

alignment_scores:
 Quality: 109.50 Length: 458
 Ratio: 0.495 Gaps: 18
 Percent Similarity: 48.253 Percent Identity: 22.926

alignment_block:
 US-09-303-518D-125/rev x T34434 ..

```

Align seg 1/1 to: T34434 from: 1 to: 2232
1340 CTTTCCTTCATAGTGTTCACGACTTTCGCAACGCGGCGGTATTC 1291
      ||||| ..... |||||
1425 ProThrGlySerThrGlnSerSerThrIleLeuSer.....Se 1437
      ||||| ..... |||||
1290 GTATTTCGCGGCGAGACGAGCTGCACAAAGCAGAGCTCTTCTGTCA 1241
      ||||| ..... |||||
1437 rThrIleSerGlySerThrGln...HisThrIleMetSerIleAlaSerS 1453
      ||||| ..... |||||
1240 ATTCACGAACCAACCAATGCTCGCGCTGCGGTATCGCGCAGATTA 1191
      ||||| ..... |||||
1453 ergly.....SerThrProSerThrAsn 1461
1190 TCGCGCAAAAGCAGGCTGGCGAGATATCCAAAGGCGATCAGCGCTG 1141
      ||||| ..... |||||
1462 SerGln.....ThrGlySerThr 1467
1140 AGTACCAATCGCACCATGCGCGGTGCGCGCTTGAAGCGCTGTGTA 1091
      ||||| ..... |||||
1467 rValThrMetGlySerSerThrSerGlyValSerThr..... 1480
1090 ACTTGAAGATTGTTTTCAGGAATGCGCGAGGTTGTACGCGTGA 1041
      ||||| ..... |||||
1480 ..... 1480
1040 GAGTATTTCGCGGCTGCGCGGCGCAACCGCAACGCTCTTGTCTGCG 991
      ||||| ..... |||||
1481 .....SerSerAlaSerThrGlnProGlnMetSerThrSergl 1494
990 GCGTTCCTTCGATACGAATCTGATTGCGTACGCTCCCAATTAATGT 941
      ||||| ..... |||||
1494 nGlySerSerAlaGlySer.....ThrV 1502
940 GCGCGCTTGTGTAATCGCGCG.....TTCATACCGAACCGGAATC 897
      ||||| ..... |||||
1502 aAlaSerSerThrAlaSerProAlaAlaSerSerThrAlaProSerSer 1518
896 ACGCGTTCCTGCTGTCAACCAATCGCGCGCAATTA.....ATTG 856
      ||||| ..... |||||
1519 ThrGlyThrMetSerSerThrSerSerGlyThrValGlySerThrIleSe 1535
855 CGATCTTTCGACCCCAACGATGACGACGAGCGGCTTGTGACTT 806
      ||||| ..... |||||
1535 rGlnSerSerThrThrAlaSerAlaSerSerGlnThrGly...SerThrV 1551
805 GAGAACCACTAGGCGCATTCACGCGCTGCTGCAGACGCGCTGTGCA 756
      ||||| ..... |||||
1551 alThrMetGlySerSerSerThrSerGlyValSerThrSerSerAlaSer 1567
755 AACCAACGCGCAATGTAATTAATCACTTGAATTAATTAATGATGTC..CACAC 709
      ||||| ..... |||||
1568 SerThrGlnProGlnMetSerThrSerGlnGlySerSerAlaGlySerTh 1584
708 GGTTTTATTCGCGCGACCGGCTCGATGAATGAATGCGGTGCGCACTCA 659
      ||||| ..... |||||
1584 rValAlaSerSerThrAlaGlyLeuValSerThrSerThrValProSerS 1601
658 AACCGGCGATGCGGCGCGCGCAATTCATGCTGTTGATGCTGGCAGCA 609
      ||||| ..... |||||
1601 erThrGlyThrMetGlySerThrSerSerGlyThrValGlySerThrIle 1617
608 TTTTCAGACGCGACGCTGCGCGCATGCTTACAAACATGATTTG... 561
      ||||| ..... |||||
1618 SerGlnSerSerThrThrAlaSerAlaSerSerGlnThrGlySerThrVa 1634
560 .....CGTTCGTCAAACGCGCTCAATA 539
1634 lThrMetGlySerSerSerThrSerGlyValSerThrSerSerAlaSerS 1651
538 CCAACAGCGCGGCTTGAATCTCTG...GCGGCTCTTGAATATGACAC 492
      ||||| ..... |||||
1651 erThrGlnProGlnMetSerThrSerGlnGlySerSerAlaGlySerThr 1667

```

```

491 GTAGGTCGCGACGACGCGATTGGTCCATTCGATTCAGCAAGATGCC 442
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1668 ValAlaSerSerThrThrGlyLeuValSer.....ThrSerThrVa 1681
441 GAAGGCGCGCGCATCGACGCGCAAGATTTCCTGACGCGAGCGGTGCCA 392
      : :|||:|||||:|||||:|||||:|||||:|||||:|||||:
1681 LProSerSerThrGlyThrMetGlySerThrSerSerGlyThrValGlyS 1698
391 GCGCAGTCACAAACCGGATTGGATC.....AGTTTGGCGCGCATCTTC 348
      |||:|||||:|||||:|||||:|||||:|||||:|||||:
1698 erThrIleSerGluSerSerThrAlaIleAlaSerAlaSerSerGlnThrGly 1714
347 TCGCGCGTTAAGTTT.....GCCAGCGCTTCAGTGGCTAGCGCTCAAA 304
      ||| :|||:|||||:|||||:|||||:|||||:|||||:
1715 SerThrValThrMetGlySerSerSerThrSerGlyValSerThrSer 1731
303 CTCGATTTCGTGCTGGCT.....TCACGGGCA 275
      |||:|||||:|||||:|||||:|||||:|||||:|||||:
1731 TAlaSerSerGlyGlnProGlnMetSerThrSerGlnGlySerSerAlaG 1748
274 TCACGCTGACTGAGTACGCGCTTTTTCGCCCGTGAATCGCGCGATT 225
      |||:|||||:|||||:|||||:|||||:|||||:|||||:
1748 LysThrThrValValSerSerThrAlaIleSerProAlaIleSerSerThrAla 1764
224 TTGCCTGAAGCGCGCGCAGTAAACACACGCC.....GGATT 187
      |||:|||||:|||||:|||||:|||||:|||||:|||||:
1765 ProSerSerThrGlyThrMetSerSerThrSerSerGlyThrValGlySe 1781
186 CTTTTCGCTTCGAACAGCAGCTTGGCTTTTTCAGCGCATCGCTTCCT 137
      |||:|||||:|||||:|||||:|||||:|||||:|||||:
1781 rThrMetSerGlnSerSerThrAlaIleSerThrThrSerThrIsthrGlyS 1798
136 TGACTTTCATCGAGGCGCGCATACGCGCATATTCCTGCCAAGCAAGCG 87
      ||| :|||:|||||:|||||:|||||:|||||:|||||:
1798 erThrValThrLeugly.....SerSerSerThrSerSerAsnGln 1811
86 ACTTGCGTAAATGCGCGGCGCTCG 63
      |||:|||||:|||||:|||||:|||||:|||||:|||||:
1812 MetSerThrSerGlnGlySerSer 1819
seq_name: p1r2.J00405

```

```

seq_documentation_block:
  hypothetical 119.5K protein (uvra region) - Micrococcus luteus
  N:Alternate names: ORF 1 protein
  C:Species: Micrococcus luteus, Micrococcus lysodeikticus
  C:date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 04-Feb-2000
  C:Accession: J00405
  R:Shiota, S.; Nakayama, H.
  M:1. Gen. Genet. 217, 332-340, 1989
  A:Title: Micrococcus luteus homolog of the Escherichia coli uvra gene: identification of
  A:Reference number: S04781; MUID:89364717
  A:Accession: J00405
  A:Molecule type: DNA
  A:Residues: 1-1106 <SHI>
  A:Cross-references: EMBL.X15867
  A:Note: This reading frame extends between two stop codons and does not begin with a sta
  C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology

```

```

alignment_scores:
  Quality: 108.50      Length: 486
  Ratio: 0.583        Gaps: 25
  Percent Similarity: 38.272      Percent Identity: 23.868

```

alignment_block:

US-09-303-518d-125 x J00405 ..

Align seg 1/1 to: J00405 from: 1 to: 1106

18 AGGTCTAAACCTGCGCATTCGCGGCGACAGCCGCAAGCCGTTTACGAGC 67

```

|||||:|||||:|||||:|||||:|||||:|||||:|||||:
528 ArgGlyArgProValThrValIsthrGlyArgIsthrProProAlaIsthrAlaArgI 544
68 GCGCGCGCATTCACGCAAGTCGGCTTCCTGGCGCAAGATATCCGGTATG 117
      ||| :|||:|||||:|||||:|||||:|||||:|||||:
544 YstIsthrValIsthrLeuGlyAlaIsthr..... 553
118 CGGCCCTCGATGAAGTCAAGGAGCGCATGCCGTCAAAAAAGCCCAAGT 167
554 .....AspGlyProGlyAlaIsthrAlaIsthrArgGlyProGly 564
168 GCTGTTTGAACAAAAAGATCCGGCGCTGGTCTTACTGCGCGGCTT 217
      ||| :|||:|||||:|||||:|||||:|||||:|||||:
565 Ala.....GlnGlyAspProGlyProAlaGlyValProAlaGly.. 577
218 CAGCAAAATCGCGCGGATTACCGGTGGGAAAAAGCGCTACTTCAGTCA 267
577 ..... 577
268 GTGCTATTGCGCTGAGGCAAGCAAGCAAAATGAGTTGAACGCTACGC 317
      |||:|||||:|||||:|||||:|||||:|||||:|||||:
578 .....ArgArgProArgValProGlnProArg 586
318 ACCTGAAGCGCTGCAAAATTTAAGCGCGCAAGAGTGGCGCGCACTGA 367
      |||:|||||:|||||:|||||:|||||:|||||:|||||:
587 .....AlaIsthrGlyArgIsthrProLeuArgArgGlyPro..... 598
368 TCCATTCGGTTGTGACTGCGCTGCGCGCGCGCTCCGTTACACCAAAAT 417
599 .....AlaIsthrProGlyIsthrIsthrAsp 607
418 CTTGCGCTGATGCGCGCGCTTCGCCATCTTCGTAATGC..... 458
607 rGlnIsthrAlaGlyArgArgProLeuArgProArgArgIsthrIsthrAsp 623
459 .....GATGCAACCAATCCG 475
624 ProAlaProAlaGlyGlnProProIsthrArgAspProProAlaProAl 640
476 TGGCTGCGCGACCCACGCGTATATCAAGAGCGCGCGGATTCGAA 525
      ||| :|||:|||||:|||||:|||||:|||||:|||||:
640 agIsthrProArgIsthrProIsthrArgArgAlaIsthrGlyIsthrAsp 657
526 CGCGCGCTGTGTGATAGCGCTTTCAGCGCAAGCAAGCAAAATCCATGTTG 575
      ||| :|||:|||||:|||||:|||||:|||||:|||||:
657 rGlyArg.....GlyGlyLeuAspArgGlyIsthrArgProSerArg 669
576 TAAGCAAGCTGGCGCAAGCT..... 596
670 ...GlyArgValArgArgArgGlyArgAlaLeuGlyLeuProGlyIsthr 685
597 .....GCCGTGCAAAATGCTGCAACATCGCAACACATGAA 633
685 rGlnGlyGlnIsthrAlaVal.....ArgIsthrArgArgLeuProI 698
634 TTGCGCGCGCGCGCATCTTCGCGGTTTGTAGTGGCAGCAGCATTCATTTCAT 683
      |||:|||||:|||||:|||||:|||||:|||||:|||||:
698 euArgProPro.....LeuIsthr 703
664 CGA...GCCGTGCGCGCGCAATTAACCGTGTGACCATCATTAATGACAG 730
      |||:|||||:|||||:|||||:|||||:|||||:|||||:
704 ArgGlyAlaGly..... 707
731 ATGTAAATACCATTCGCGCTTGTGTGCAACAGCGCGCTGCAACAGCAG 780
      |||:|||||:|||||:|||||:|||||:|||||:|||||:
708 .....AlaIsthrArgProGlyIsthrAla 717
781 CGCGTATTCGCTAGTGTGTTCTCAAGTCAACAAACCGCGCTTCGCG 830
      |||:|||||:|||||:|||||:|||||:|||||:|||||:
717 rGlyAlaAspGlyProArgArgProGlyGlnGlnProGly..... 729
831 TACCGTTTGGTGGCAAGATATGCAAAATTACGCGGCGCGAAT..... 875
      |||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

730 .....GlyArgLeuGlyPr 734
876 .GGTTACACAGACAAACCG.....CGTGAATTCGGCTTGGTATGTAAC 918
734 oGlyProAlaArgGlyProHisGlyArgAspGlyArgValArgLeuArg 751
919 GGGCGGATTAACAAGGCGGACGATTTATTTGGAGCGGTACACAAATCA 968
751 InvAlaHisAlaAspGlnArgAspProLeuGlnGlyPro...GlyGlnPro 766
969 GATTCGCGTATATCAAGAGAGAGCGGACAAAGAGCTGTCGGCTGGGTG 1018
767 AlaGlnArgGlyGlnAlaArgAlaArgProAlaArgValArgGlyGly 783
1019 GCGCGACGCGGACAAATACATCTCAATCAGCGGTACAAACCTCGCCATTTC 1068
783 rArgAlaProGlyGln.....GlyGlyProArgGlyProGlnProHis 798
1069 CTGAACAAACAACTCTCAAGTTCA.....CACAGCGGTCAAGCGGG 1112
798 rGAlaHisAlaThrLeuGlnProArgHisLeuHisGlyArgValArg 814
1113 GCAGCGCGCATGTCGCGATGCTTACTTACAGAGCG.....CGTGA 1153
815 AspProGln.....AlaLeuArgGlyAspProArgGlyGly 826
1154 TGCCCTGGATATCTGCGCACCGCTTGTTCGCGGATTTATGTCGGC 1203
826 uGlyProGlyLeuProAlaGlyProValLeu..... 837
1204 GATACGACAGCGGACGCGATGCTTGGTTCGTTGCAATTCGACGAAGA 1253
838 .....GlnHisGlnGlyArgAlaLeuArgGlyValArgGlyArg 851
1254 CCTGGCTTTGTCAGCTTCCTGTCGCC...GGGCAATACGA...ATACG 1297
852 HisAlaGlnAspArgAspGlyLeuProAlaGlyArgLeuArgAlaVal 868
1298 G.....CCGCGTGTGGCGAAAGCGTCGGAACCATTTGAG 1332
868 gGlyValProArgGlyProValGlnProGlyAspAlaArgGlyHisLeu 885
1333 AAGGAAG 1340
885 InGlyGln 887
seq_name: p1r2:A98048
seq_documentation_block:
H+/K+-exchanging ATPase (EC 3.6.1.36) [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
C:Accession: A98048
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; M01D:21429245; PMID:11544234
A:Accession: A98048
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-898 <RUP>
A:Cross-references: GB:A007317; PIDN:AL00214.1; PID:G15459062; GSPDB:GN00174
C:Genetics:
A:Gene: PacI
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
C:Keywords: hydrolase

```

alignment_scores: Quality: 108.00

Length: 356

Ratio: 0.617 Gaps: 15
Percent Similarity: 49.157 Percent Identity: 23.315
alignment_block:
US-09-303-518D-125 x A98048 ..
Align seg 1/1 to: A98048 from: 1 to: 898

```

28 CTGCCCATCCGGGCGACAGCGGAGCAAGCCGTTTACGAGCGCCGCGCAT 77
111 .....GlyArgLeuGlyProAlaArgGlyProHisGlyArgValArg 299
287 LeuAlaValAlaAlaLeuProGlnGlyLeu.....ProAlaIle 299
78 TACCGAAGTCGCGTTC.....CTTGGCAAGAAATATGCGGATGCGCCCT 124
299 eValThrIleValLeuSerLeuGlyThrGlnValLeuAlaIleValSAlaHis 316
125 CGATG.....AAGTCAAAGGAAGGCGATGCGGTCAAAAAGCGCAATG 168
316 eXlleValArgGlyLeuProAlaValAlaGlnThrLeuGlySerThrGluile 332
169 CTGTTGAAGACAAAAGAAAT.....CC 191
333 IleAlaSerAspIleThrGlyThrLeuThrMetAlaMetThrValGln 349
349 uLysValPheTyrAspAlaValLeuHisAspSerValAlaValLeuAlaAspThr 366
242 GTGGC...GAAAAGCGGTACTTCACTGATGCTGATGCTGCGTGAAGGC 288
366 euGlyLeuGlnMetProLeuArgSerValValLeuAlaAspThr 382
289 AACGACGAATGAGTTTGAACGCTACGACCTGAAGCGGTGCAACTT 338
383 LysIleAspValGln.....GlyAsnIle 390
339 AAGCGGCAAGAGTGGCGCGCAACCTGATCCAAATCGGTTGTGGAGCTG 388
390 uLleGlyAspProThrGlnThrAlaPheIleGlnThrAlaLeuAspLysG 407
389 CGCTGCGGACCGGTCGCTTC...ACCAAAATTCCTGCGGTGATGCCAG 435
407 LYTyrAspValLysGlyPheLeuGlnLysTyrProArgValAlaGlnLeu 423
436 CGGTGCGCATCTTGTCAATGCGATGAGCAC...AATCGCGCGTGC 482
424 ProPheAspSerAspArgLysLeuMetSerThrValHisProLeuProAs 440
483 GCACCTTACGGTCAATTATCAAGAAAGCCGCGAGGATTTCAACGCGGCC 532
440 pSerArgPheLeuValAlaValLysGlyAlaProAsp..... 452
533 TGTGTGATTTAGCGGTTTGACGCAAGCAAAATCCATGTTTGAAGCA 582
453 .....GlnLeuLeuLysArgCysLeuLeuArgAspLysAla 464
583 GCTGGCGGACAGCTGCGGTGAAATAGTGCAC...ATCGAAACACA 629
465 GlyAspIleAlaProIleAspGlyLysValThrAsnLeuIleHisThrAs 481
630 TGAATTGGCGCGCCGCAATCTCGC.....GGTTGAGTGGCAGCGACA 673
481 nAsnSerGlnMetAlaHisGlnAlaLeuArgValLeuAlaGlyAlaLyr. 497
674 TTCAATTCATCGACCGCGTGGCGCGAATAAACCGTGTGACCAATCAAT 723
498 .....LysIleIleAspSerIleProGlnAsnLeuThrSerGlnGluLeuGln 513
724 TATCAAGATTAATTACATGTCGCGGTTGTTTGAACAGCGCGCTGAA 773
514 ...AsnAspLeuIlePheThrGlyLeuIle.....GlyMetIleAs 526
774 CACCGAGCGCGTATTCGCTAGTGTTCCTCAAGTCAACAAACCGCGCC 823

```


